Genetic Diversity and Population Structure of the critically endangered freshwater fish species, the Clanwilliam sandfish (*Labeo seeberi*)

By

Shaun Francois Lesch

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in the Faculty of Natural Science at Stellenbosch University

Supervisor: Dr C. Rhode

Co-supervisor: Dr R. Slabbert

Department of Genetics

December 2020

Declaration:

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

. . . . . .

Date: December 2020

Abstract:

Labeo spp. are large freshwater fish found throughout southern Asia, the Middle East and Africa. The genus is characterised by specialised structures around the mouth and lips making it adapted to herbivorous feeding (algae and detritus). Clanwilliam sandfish (*Labeo seeberi*) was once widespread throughout its natural habitat (Olifants-Doring River system), but significant decreases in population size have seen them become absent in the Olifants River and retreat to the headwaters in the tributaries of the Doring River. Currently sandfish are confined to three populations namely the Oorlogskloof Nature Reserve (OKNR), Rietkuil (Riet) and Bos, with OKNR being the largest of the three and deemed the species sanctuary. Sandfish play an important role cycling nutrients and maintaining algae levels in aquatic ecosystems and is therefore an important species for conservation in the Cape Floristic Region (CFR) as it maintains river health. This thesis contributes toward the establishment of an effective Biodiversity Management Plan (BMP-S) for Clanwilliam sandfish. Phylogenetic analysis of *Labeo* spp. using two mitochondrial DNA regions (mtDNA), Cytochrome oxidase subunit 1 (*CO1*) and Cytochrome b (*Cytb*), showed that *Labeo seeberi* was most closely related to *L. vulgaris* (also known as *Labeo niloticus*). Phylogenetic analysis also recovered the *Labeo niloticus* group (LNG), *Labeo forskalii* group (LFG) and *Labeo coubie* group (LCG) as proposed by Reid 1985 and Ramoejane *et al*. 2016. Extrapolating from Ramoejane *et al*. 2016, *L. seeberi* is part of the *Labeo umbratus* group (LUG) and therefore its closest relative is *Labeo capensis* (geographically its closest relative as well). Using *L. capensis* as reference, it is postulated that *L. seeberi* reaches sexual maturity at ± 4 years of age (250mm TL) and grows at 40-60mm per year up to six years where after the growth rate decreases steadily. Population genetic studies using microsatellite markers and mtDNA (*D-Loop*) revealed no genetic differentiation between the three populations (OKNR, Bos and Riet) and no sign of significant inbreeding (F_{IS}) or relatedness (r) indicating gene flow maintaining genetic diversity. Effective population size (N_e) of OKNR was as expected much higher than Riet and Bos. Genetic evidence thus corroborates the assumption that the OKNR is the main breeding population that then migrate to Riet and Bos maintaining gene flow and genetic diversity. Thus, the collective of OKNR, Riet and Bos must be handled as a single Evolutionary Significant Unit (ESU), with OKNR and Riet-Bos being separate Management Units.

Opsomming:

Labeo spp. is groot varswatervisse wat in suider Asia, die Midde-Ooste en Afrika voorkom. Die genus word gekenmerk deur die gespesialiseerde mond en lip strukture wat spesifiek aangepas is tot ’n herbivoriese dieet (alge en detritus). In die verlede was Clanwilliam sandvis (*Labeo seeberi*) wyd verspreid oor sy natuurlike habitat (Olifants-Doring Rivierstelsel), maar ’n noemenswaardige afname in populasiegrootte het daartoe gelei dat sandvis tans afwesig is in die Olifantsrivier en slegs klein populasies in die sytakke van die Doring Rivier voortbestaan. Tans word sandvis tot drie populasies beperk, naamlik die Oorlogskloof Natuurreservaat populasie (OKNR), die Rietkuil (Riet) populasie en die Bos populasie. Die OKNR populasie is heelwat groter as die ander twee en word beskou as ’n bewarea vir die spesie. Sandvis speel ’n belangrike rol in varswater ekosisteme, deur alge-vlakke en die sirkulering van voedingstowwe te handhaaf. Daarom is dit belangrik om sandvis te bewaar siende dat dit die riviere van die Kaapse Blomme Streek (KBS) se gesondheid handhaaf. Hierdie tesis poog om by te dra tot die vestiging van ’n effektiewe Biodiversiteitsbetuursplan (BMP-S) vir die Clanwilliam sandvis. Met die gebruik van twee mitochondriale DNA-streke (mtDNA), Cytochrome oxidase subunit 1 (*CO1*) en Cytochrome b (*Cytb*), wys die filogenetiese analise dat sandvis naaste verwant is aan *Labeo vulgaris* (ook bekend as *Labeo niloticus* of Nile carp). Die filogenetiese analise identifiseer ook die *Labeo niloticus* groep (LNG), *Labeo forskalii* groep (LFG) en die *Labeo coubie* groep (LCG) soos voorgestel deur Reid 1985 en Ramoejane *et al*. 2016. Aflei vanuit hierdie groepe, voorgestel deur Ramoejane, 2016, plaas dit die sandvis in die *Labeo Umbratus* groep (LUG) en is *Labeo capensis,* dus die mees naverwante spesie (sandvis en *L. capensis* is ook geografies naaste aan mekaar). Deur *L. capenis* te gebruik as verwysing kan daar gepostuleer word dat *L. seeberi* seksuele volwassenheid bereik teen ±4 jaar oud (250mm TL) en dat dit teen ’n tempo van 40-60mm per jaar groei vir die eerste ses jaar, waarna dit stelselmatig verminder. Populasie-genetiese studies, met die gebruik van beide mikrosatelliet-merkers en mtDNA (*D-loop*), identifiseer dat daar geen noemenswaardige genetiese differensiasie tussen die drie populasies (OKNR, Riet en Bos) is nie, asook geen noemenswaardige inteling (F_{IS}) of verwantskap (r) nie. Dit dui daarop dat daar geenvloei tussen die populasies is en sodoende die genetiese diversiteit handhaaf. Die effektiewe populasie grootte (N_e) was na verwagting heelwat groter vir die OKNR as vir Riet en Bos populasies. Genetiese bewyse steun dus die aanname dat die OKNR die hoof broeipopulasie is en dat individue dan migreer na Riet en Bos, wat sodoende geenvloei en

genetiese diversiteit onderhou. Die OKNR-, Riet- en Bos-populasies moet dus as een enkele Evolusionêre betekenisvolle eenheid (EBE) beskou word met twee afsonderlike Bestuurseenheid (BE), OKNR en Riet-Bos.

Acknowledgements:

I would like to extend my gratitude to CapeNature and the University of Stellenbosch for the funding and the use of their facilities throughout the course of this project. Also, a big thank you to Dr Martine Jordaan and her team for the collection of biological samples. To my supervisors Dr R. Slabbert and Dr C. Rhode, an immense thank you not only for your scientific advice and guidance, but also for your immeasurable patience, diligence and support. Lastly to my supervisors and friends, thank you for carrying me through this adventure.

Table of Contents:

List of Figures:

List of Tables:

List of Abbreviations

CFR	Cape Floristic Region
IUCN	International Union for Conservation of Nature
SAIAB	South African Institute for Aquatic Biodiversity
OKNR	Oorlogskloof Nature Reserve
DWA	Department of Water Affairs
BMP-S	Biodiversity Management Plan for Species
AOO	Area of Occupancy
DENC	Department of Environment and Nature Conservation
WUAs	Water User Associations
N_e	Effective population size
DNA	Deoxyribonucleic Acid
mtDNA	Mitochondrial Deoxyribonucleic Acid
CTAB	Cetyltrimethylammonium Bromide [$(C_{16}H_{33})N(CH_3)_3Br$]
Cytb	Cytochrome b
CO1	Cytochrome c Oxidase subunit 1
D-loop	Mitochondrial control region
µM	Micromole
µl	Microlitre
ml	Millilitre
ng	Nanograms
PCR	Polymerase Chain Reaction
sec	Seconds
min	Minutes

°C	Degrees Celsius
P-value	Probability value (as a statistically significant limit)
H_O	Observed Heterozygosity
H_{Enb}	Expected Heterozygosity
A_N	Number of alleles
A_E	Effective number of alleles
A_R	Allelic richness
A_{PR}	Allelic private Richness
N_{ind}	Number of individuals
Fr_{null}	Null allele frequency
K2P	Kimura 2-Parameter model
bp	Basepair
F_{ST}	Wright's Fixation Index (subpopulation relative to the total population)
F_{IS}	Wright's Fixation Index (individual relative to the sub-population, equal to the inbreeding coefficient - f)
F_{SC}	Derivative of Wright's Fixation Index adapted for hierarchical AMOVA (sub-population relative to the group of populations)
F_{CT}	Derivative of Wright's Fixation Index adapted for hierarchical AMOVA (group of populations relative to the total population)
HWE	Hardy-Weinberg Equilibrium
FCA	Factorial Correspondence Analysis
AMOVA	Analysis of Molecular Variance
r	Relatedness
IAM	Infinite Alleles Model

SMM	Stepwise Mutation Model
TPM	Two-Phased Model
MUSCLE	Multiple Sequence Comparison by Log-Expectation
S	Number of polymorphic sites
h	Haplotype diversity
π	Nucleotide diversity
Φ_{ST}	Pairwise molecular differentiation
N_h	Number of haplotypes
K	Number of clusters

Chapter 1: Literature Review:

1.1 Introduction:

Freshwater fishes are some of the most threatened organisms on the planet (Carizzo *et al*., 2013), largely due to habitat degradation, water flow modification and the introduction of alien fish species (Dudgeon *et al*., 2006; Gene, 2007; Leprieur *et al*., 2009). These factors have led to the decline of global freshwater fish biodiversity (Leidy and Moyle, 1998; Pauly and Zeller, 2016).

An assessment of threats to the southern African aquatic ecosystems found South Africa to be no exception to this trend, with invasive species, water abstraction and water flow modification listed as the major causes (Darwall *et al*., 2009). Of the 355 southern African freshwater fish species that were assessed, 12 species were critically endangered, 19 were endangered, 9 were vulnerable, 9 were near threatened, 235 were of least concern and 71 were data deficient. Of the 12 species that were ranked as critically endangered, one was *Labeo seeberi* (Darwall *et al*., 2009; Ramoejane, 2016). The *L. seeberi* evaluation was based on severe declines in population sizes resulting from predation by non-native fishes (such as small mouth bass, spotted bass and bluegill sunfish), as well as habitat degradation (Lubbe *et al*., 2015). The conservation of the other 11 southern African *Labeo* species were evaluated as being of least concern, but were still facing the same threats as *L. seeberi* (Darwall *et al*., 2008). This is in large part due to their greater distribution range, greater numbers, fewer instream barriers and less drastic water level fluctuations between seasons (Darwall *et al*., 2008). *Labeo* spp. are large herbivorous fish that are important organismal components of aquatic ecosystems and are a high conservation priority in South Africa (Ramoejane, 2016).

1.2 Cape Floristic Region:

Located at the south-western tip of South Africa, the Cape Floristic Region (CFR) stretches from the Cederberg in the north-west, around the Western Cape coast and into the Eastern Cape up to the Nelson Mandela Metropole (Linder *et al*., 2010; De Moor and Day, 2013).

World famous for its dramatic and varied land- and seascapes and its astonishing diversity of plant and animal life, it is globally recognised as a biodiversity hotspot. This is highlighted by the regions freshwater fish variety, serving as host to 24 indigenous species, 17 of which are endemic to the region (Skelton, 2001; Linder *et al*., 2010; Chakona and Swartz, 2013; De Moor and Day, 2013; Weyl *et al*., 2014). The CFR's biodiversity stems from the complex geological and climatic history of this region, such as extensive uplifting and mountain building, major sea level changes (regressions and transgressions) and periods of either wet or dry conditions resulting in a very diverse landscape and freshwater fish endemism due to the geographic isolation in individual river systems (Skelton, 1994; Swartz *et al*., 2008; Linder *et al*., 2010; Skelton and Swartz, 2011; Chackona *et al*., 2013; De Moor and Day, 2013).

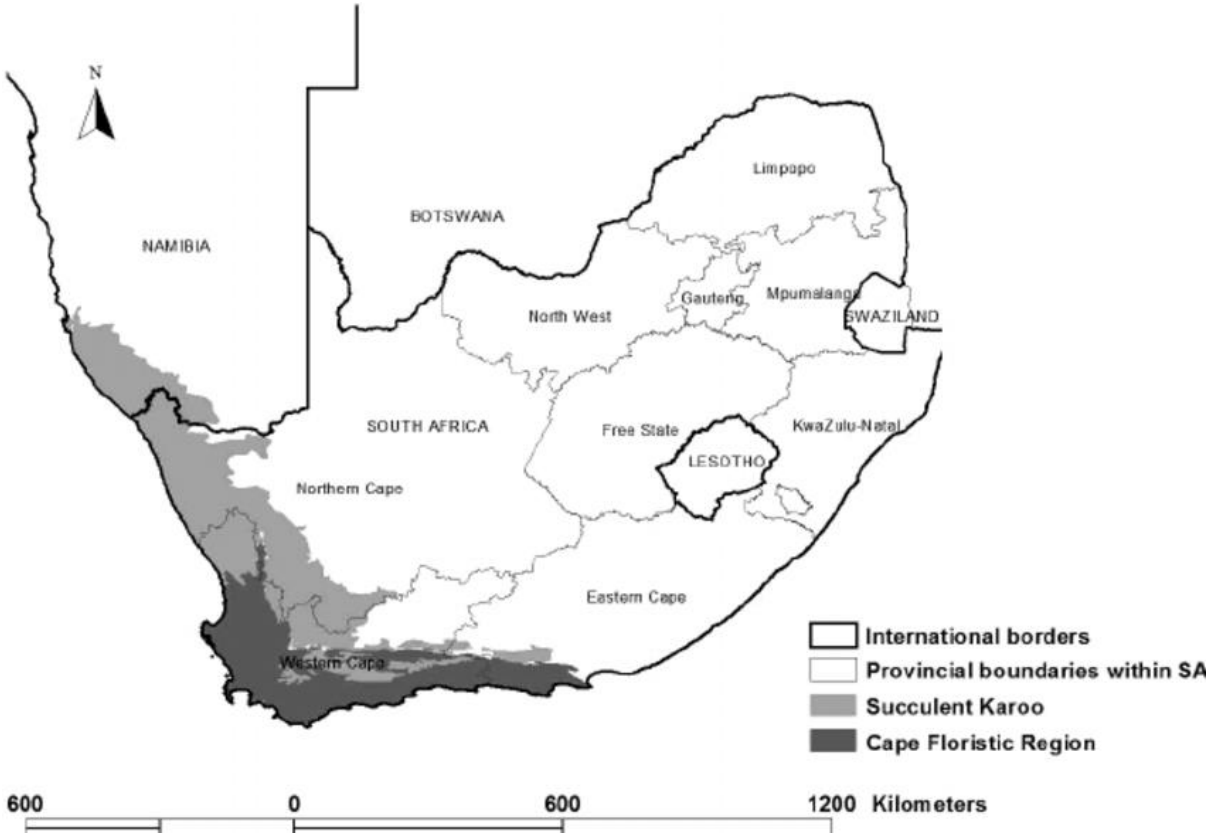


Figure 1.1: The Cape Floristic Region and The Succulent Karoo with regards to their geographical location within southern Africa. (taken from Brownlie *et al*., 2005)

Nearly all of the 24 described indigenous fish species to the CFR, are on the International Union for Conservation of Nature (IUCN) Red List; with three species being classified as vulnerable, 10 species as endangered and four species as critically endangered (IUCN 2013). The threats involved in the decline of these endemic freshwater fish species include habitat loss and fragmentation, hydrological alteration, climate change, overfishing, pollution, and

predation by and competition with alien invasive fish (Tweddle *et al*., 2009; Chackona and Swartz, 2012). Currently, 16 alien invasive freshwater fish species have established self-sustaining populations in the rivers of the CFR (Marr, 2012), of which a number of species can be linked to the decline in native fish populations (De Moor and Bruton, 1988; Tweddle *et al*., 2009). The pressures of alien invasive predatory fish, combined with the stresses of habitat degradation, has resulted in the absence of many native fish species in the lower reaches of tributaries and main stream rivers in the CFR (Tweddle *et al*., 2009; De Moor and Day, 2013; Weyl *et al*., 2013). This means that native fish populations have become even more highly fragmented, with many species now largely confined to the headwater reaches of streams (Swartz *et al*., 2004; Tweddle *et al*., 2009; Chackona and Swartz, 2012). Conservation and management of this region in order to maintain the diversity and well-being of species is thus of increasing importance (Paxton *et al*., 2012).

1.3 *Labeo seeberi*:

1.3.1 Biology, Habitat and Ecology:

Labeo seeberi (Figure 1.2) or more commonly known as the Clanwilliam sandfish is one of the larger *Labeo* species, with rare individuals having weighed more than 2kg and measured as much as 650mm in length. These recordings are, however of the extremes, whereas the modal length for these fish in the mainstream Doring River is approximately 500mm (Gaigher, 1973; Paxton *et al*., 2002; Paxton *et al*., 2012). Individuals from tributary populations are however growth limited as a result of food and space, rarely exceeding 250mm in maximum length. Thus making them susceptible to falling within the prey size range for predatory invasive fish (Paxton *et al*., 2002; Paxton *et al*., 2012). *Labeo seeberi* is easily identifiable by its olive-grey skin colour, small eyes, minute scales, spindle shaped body and its most discernible feature, its well-developed papillose lips. It is also these traits that make Clanwilliam sandfish adapted to its benthic feeding, scraping algae, diatoms and detritus from the rocky, river bottom using its sucker-like mouth (van Rensburg, 1966; Skelton, 1987; Skelton, 2001; Paxton *et al*., 2012).

Figure.1.2: Picture of an adult *Labeo seeberi* specimen taken during the 2013 sampling event. (Photo curtesy of Dr. M. Jordaan, SAIAB)

Clanwilliam sandfish are reported to be rheophilic, meaning that individuals seek pools or deep runs of larger rivers for feeding, overwintering and oversummering, whilst during spawning are required to travel upstream to the fast-flowing headwaters of the tributaries (Paxton, 2002). This mass upstream migration, paired with spawning takes place during spring (September – November) (Harrison, 1977; Paxton *et al*., 2012). Sexual maturity is reached once the individual reaches ±250mm in length, with older larger captive females yielding ±80 000 eggs (Jubb, 1967; Gaigher, 1973; Impson, 1997; Paxton *et al*., 2012)

There is strong evidence (although inferred from close relatives) that spawning is closely linked to rainfall and the subsequent increase in flow rate of the headwaters, bringing with it rich nutrients (Lubbe *et al*., 2015). Therefore, poor rainfall or the blocking of water-flow can lead to poor nutrient concentrations in the water downstream. Females respond to the substandard nutrient concentration and retain their eggs until, conditions are optimal or reabsorb gonads altogether if conditions do not improve (Paxton *et al*., 2012). This can result in certain years having very low recruitment success, as females refrain from spawning for that season (Gaigher, 1984; Tómasson *et al*., 1984; Potts *et al*., 2005).

1.3.2 Distribution:

Clanwilliam sandfish are geographically confined to the Olifants-Doring River system (Figure 1.3) in the Northern- and Western Cape provinces of South Africa (Skelton, 2001). The species was once widespread throughout the Olifants-Doring River system as highlighted by Harrison (1963) who in 1938 observed large aggregations of juvenile sandfish near Keerom in the upper reaches of the Olifants River. He also reports on having witnessed thousands of sandfish amassed below/downstream of dam walls of the Bulshoek and Clanwilliam Dams (in the middle reaches of the Olifants River), during the annual September spring spawning run (van Rensburg, 1966; Harrison, 1977). Clanwilliam sandfish had last been recorded in these middle reaches (Bulshoek and Clanwilliam dam) of the Olifants River in 1958 (Paxton *et al*., 2012; SAIAB Database). Further evidence suggest that they have been extirpated from the Olifants River as a whole, as no specimens have been recorded since 1987 (Lubbe *et al*., 2015). Currently, the sandfish population are confined to the middle and northern reaches of the Doring River and its isolated tributaries namely; Oorlogskloof-Koebee, Gif, Kransgat, Biedouw, Tra-Tra and Matjies Rivers where sandfish have been recorded in the last five years (Paxton *et al*., 2012; Lubbe *et al*., 2015).

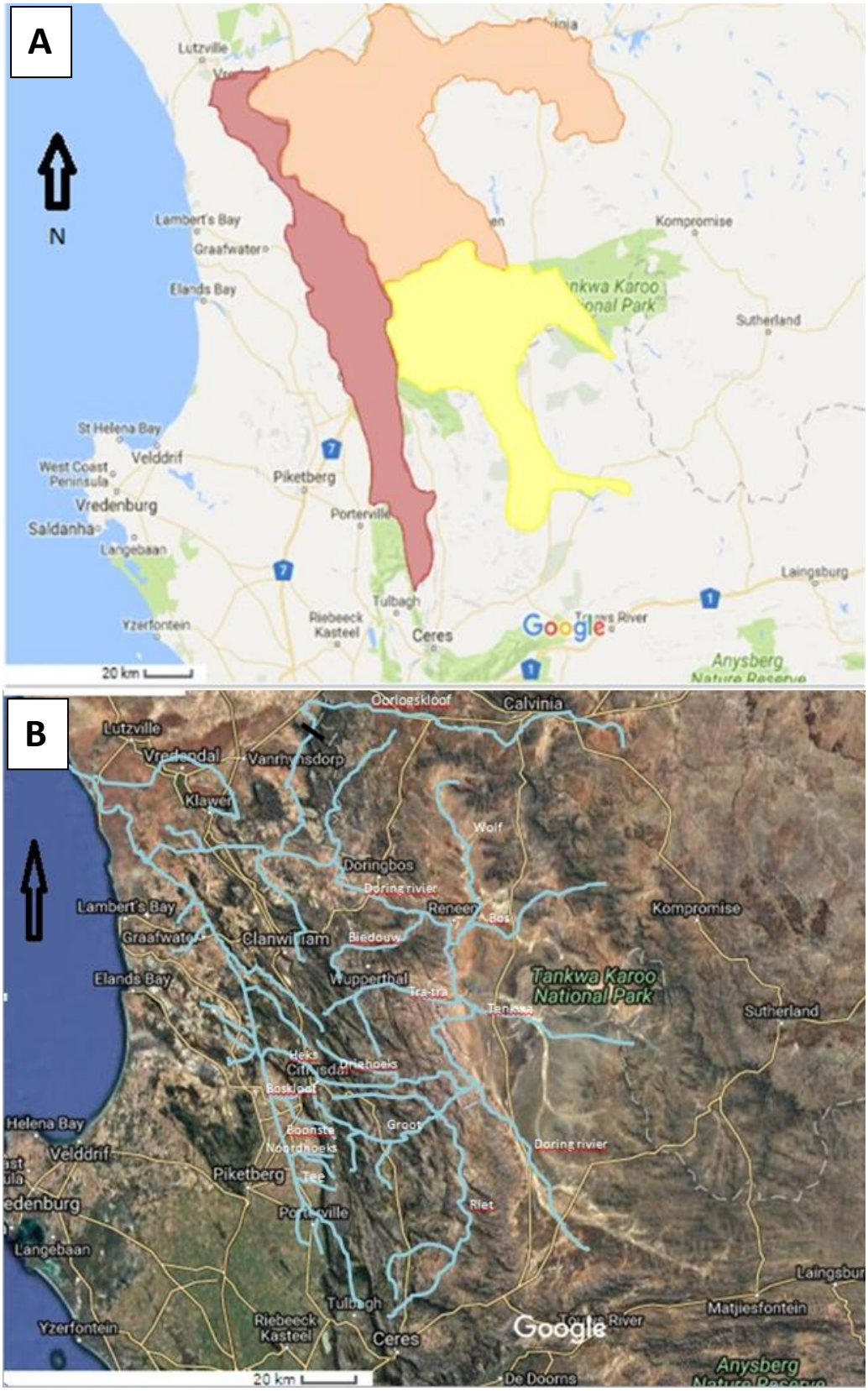


Figure 1.3: A) Map showing the historical distribution of *Labeo seeberi* in the Olifants Doring River system. The red area indicates the extirpation of *L. seeberi* in this region, whilst the yellow area signals the low frequencies of sandfish with no recruitment. The orange area indicates the current distribution of sandfish. B) Represents the rivers of the Olifants Doring River system.

Most notable of these tributaries is the Koebee River, called the Oorlogskloof in its upper reaches. The Koebee River is integral in linking the upstream migration of Clanwilliam yellowfish (*Labeobarbus capensis*), sawfin (*Barbus serra*) and Clanwilliam sandfish (*Labeo seeberi*) from the Doring River to the Oorlogskloof gorge (Impson, 1997; Abrahams and Pretorius, 2000; Ramollo *et al*., 2012). The Oorlogskloof River serves as haven for endemic and endangered fish species of the CFR, acting as a safe and successful spawning and nursing site for these fish (Impson, 1995; Ramollo *et al*., 2012). This is in part due to the relative inaccessibility of the Oorlogskloof River as it flows through the steep slopes of the Oorlogskloof gorge just south of Nieuwoudtville, a stretch of only 18.66 km of non-perennial river preventing much of the habitat destruction as seen throughout the rest of the tributaries of the Olifants-Doring River system. The Oorlogskloof Nature Reserve (OKNR) provides habitat for the only known viably recruiting subpopulation of the species, as it is the only habitat which is both free of predatory alien species and provides suitable habitat for spawning and successful recruitment of juvenile fish (Lubbe *et al*., 2015). Alien predators are restricted to the lower reaches of the Oorlogskloof-Koebee River by means of a natural barrier of huge boulders (that result in a waterfall), just south of the Oorlogskloof Nature Reserve (OKNR) (Abrahams and Pretorius, 2000; Ramollo *et al*., 2012).

Due to very low numbers of adult sandfish and high predation by alien fish species, recruitment contributions from the remainder of the catchment are not expected to be significant (Lubbe *et al*., 2015). The Oorlogskloof Nature Reserve (OKNR) is the only pristine habitat and thus constitutes the only viable population (Lubbe *et al*., 2015).

1.3.3 Population:

Clanwilliam sandfish are listed as critically endangered due to its small geographical distribution and declining numbers (Impson and Swartz, 2007; Paxton *et al*., 2012; Lubbe *et al*., 2015). A survey carried out in 2001, followed by more extensive surveys in 2003, 2011 and 2013 sampled the main stream Doring River. Although adult fish were present in the main stream Doring River, they are rare and heterogeneously distributed. This heterogeneous distribution is most likely the result of their schooling and migratory habits in relation to

environmental factors such as food availability and flow, influencing their eventual habitat selection. The rarity of these mainstream fish can be assigned to the fact that there is little to no successful recruitment of juvenile fish. This is as result of the main stream being dominated by predatory alien invasive fish such as smallmouth bass, *Micropterus dolomieu*, spotted bass, *M. punctulatus* and bluegill sunfish, *Lepomis macrochirus*. These invasive fish prey on the juveniles as discussed in section 2.1.4 *Threats* (Paxton *et al*., 2012; Lubbe *et al*., 2015). This lack in recruitment of juveniles is further corroborated by empirical data from the 2003, 2011 and 2013 surveys indicating that there was a more than 50% decrease in the number of sites at which sandfish were caught in 2011 and 2013 as compared to 2003 (Lubbe *et al*., 2015) This further emphasises the trend of population decline of sandfish. The size class data collected during the 2013 Doring River main stream survey suggests that the current sandfish population that is persisting, does so because it is predominantly comprised of old, large fish that are beyond the prey size class of predatory alien invasive species. No indigenous fish species smaller than 400 mm (i.e. no juveniles or sub adults) were recorded indicating that there is no or minimal recruitment taking place. This suggests that the sandfish population is likely decreasing and becoming more fragmented in the Doring River main stream. Surveys in 2012, 2013 and 2014 were conducted in a number of tributaries of the middle and northern reaches of the Doring River namely; Biedouw, Tra-Tra, Matjies, Kransgat, Oorlogskloof-Koebee and Gif rivers, following reports of sandfish presence (Lubbe *et al*., 2015). These tributaries all harbour populations that are confined to very limited stretches of river, with natural barriers such as boulders and waterfalls safeguarding these sandfish populations from alien invasive species (Lubbe *et al*., 2015). With the exception of Oorlogskloof, these small isolated populations consist of very few adult fish (n<10), and are essentially boxed in by predatory invasive alien species whom prey on the young. It is thus very unlikely that any of these small isolated populations make any meaningful contribution to the overall population size. The exception may be the Biedouw River, which in 2011 reported a successful spawning, the first recording outside of the OKNR in a number of generations**.** Adult sandfish are, however, common in the 18.66 km stretch of the Oorlogskloof River as stated in section 1.2.2 *Distribution*. The southern edge of this stretch is demarcated by a waterfall, acting as invasion barrier to alien invasive species, whilst the rest of the stretch is located in a ravine with steep rocky slopes making it inaccessible for livestock and agriculture. Although a fair number of adult sandfish are present downstream of the waterfall, these individuals are unable to recruit

successfully as result of the alien invasive predators present (Paxton *et al*., 2012; Lubbe *et al*., 2015).

The OKNR thus likely serves as the last annually recruiting sandfish subpopulation, therefore making it the last suitable spawning habitat for the species. The lack of suitable spawning habitat, recruitment, mortality, low densities, ageing population and heterogeneous distribution of sandfish outside of the OKNR, make the species extremely vulnerable to extinction. As there is less recruitment of juveniles and older fish are lost from the system and not replaced, population size will decrease. The Oorlogskloof Nature Reserve is thus critical in terms of the survival of the highly threatened species, emphasising the importance for the proper conservation and management of this region.

1.3.4 Threats:

The major threats faced by sandfish in the Olifants-Doring River system are similar to the threats faced by all endemic fish of this region and has been recognised and well documented over a fair time-span (Gaigher, 1973; Scott, 1982; Impson, 1997; Impson *et al*., 2000; Paxton *et al*., 2002; Woodford *et al*., 2005; Nel *et al*., 2006; Impson and Swartz, 2007; Lubbe *et al*., 2015). Like with all the other endemics in the Olifants-Doring River system, the main threat to the survival of sandfish populations is predation and competition for resources by alien invasive fish species (Impson *et al*., 2000; Paxton *et al*., 2002; Woodford *et al*., 2005). The most notable of these predating alien fish species are smallmouth bass, *Micropterus dolomieu*, largemouth bass, *Micropterus salmoides* and bluegill sunfish, *Lepomis macrochirus*. These species were introduced to the Olifants-Doring River system by the former Department of Inland Fisheries during the earlier half of the 20th century to serve as sport fish for anglers (Roth, 1952; Harrison, 1977; De Moor and Bruton, 1988; Paxton *et al*., 2002). Although these three species present a major predatory threat to sandfish and have all but replaced the endemic fish where they occur, a different threat comes from banded tilapia *Tilapia sparrmanii* in the form of competition for food. This is especially of concern, as highlighted in section 1.3.3 *Population*, banded tilapia invaded the Oorlogskloof River, above the waterfall, thus in the OKNR. This puts them in direct competition with the sandfish population in the OKNR, and although surveys done in 2000, 2010, 2013 and 2014 concluded that the sandfish

population in this region is stable, there is no assurance that it will remain this way (Paxton *et al*., 2012; Lubbe *et al*., 2015).

An additional and potentially greater threat is the recent confirmed reports that Sharptooth Catfish (*Clarias gariepinus*) are present in the Olifants-Doring system. This species has been introduced, illegally in most cases, into all four primary river systems of the Western Cape (Paxton *et al*., 2012; Lubbe *et al*., 2015). This species may pose a bigger threat than big- and smallmouth bass and bluegill sunfish due to their ability to survive and adapt to a range of environmental conditions, their ability to survive desiccation, their omnivorous feeding habits, high fecundity, fast growth rate, dispersal ability, predatory habits and large size (Paxton *et al*., 2012).

Other threats include water quantity and instream barriers. Water resources in both the Olifants and Doring Rivers are heavily exploited. Water exploitation is especially severe in the Olifants River, where water abstraction (primarily for irrigation) and flow regulation by dams and weirs have greatly altered the flow of the river. This is compounded by increased water abstraction during the hot summer and unusually low levels of rainfall during the winter (Paxton *et al*., 2012; Lubbe *et al*., 2015). This has resulted in the Olifants River being reduced to standing pools during the dry season, with the no-flow period having increased from 5% historically to 45% currently (Birkhead *et al*., 2005). These conditions give invasive fish species an competitive advantage and have duly replaced the indigenous fish in these reaches (Paxton *et al*., 2002). While water exploitation along the Doring River is not as intensive as in the Olifants River, it is projected that abstraction in the region will increase and have a major effect on the mean water level (DWAF, 2005). In addition, a number of large-scale dam options on the Doring River have been proposed (Aspoort, Melkbosrug and Melkboom) to meet the demand for increased agriculture in the region (DWAF, 2005; PGWC, 2007).

The free movement of fish in a river system are important for the dispersal of young and moving to and from feeding, breeding and overwintering areas. Evidence from the early twentieth century indicate that annual spawning migrations were interrupted by the Bullshoek Weir, reporting large quantities of indigenous fish downstream of the barrier (Harrison, 1977). The endemic species have all but disappeared from the middle and lower reaches of the Olifants River where large-scale water resource infrastructure and water abstraction has interrupted spawning migrations and degraded aquatic habitat. In contrast,

the Doring River is still mainly free flowing. A large privately owned dam on the mainstem near Brakfontein and Department of Water Affairs (DWA) gauging weir at Aspoort are however considered substantial obstacles to fish movement during critical times of the year (Paxton *et al*., 2012). The natural flow regime of a river thus heavily impacts fish recruitment (Cambray *et al*., 1997; Poff *et al*., 1997; King *et al*., 1998; Humphries *et al*., 1999; Koehn and Harrington, 2006). Although this process is not well understood in the case of the Clanwilliam sandfish, there is substantive evidence, both anecdotal and from the ecology of closely related species (Paxton *et al*., 2012), to support the contention that the species is a synchronous rheophilic spawner requiring optimal flow and temperature conditions for successful reproduction. Natural hydrological variability, together with water regulation and abstraction is therefore likely to play a major role in year-class strength (Paxton *et al*., 2012; Lubbe *et al*., 2015).

1.4 Conservation actions - Biodiversity Management Plan for Species (BMP-S):

1.4.1 The need for a BMP-S:

As stated above, the sandfish population is becoming increasingly diminished and severely fragmented. Small adult populations are restricted to the headwaters of small tributaries. Here they are protected against alien predatory fish by means of natural barriers. It is also evident that recruitment of juvenile fish, outside of the OKNR, has seized and that these populations that do subsist in the main stem rivers represent an ageing population and are becoming more heterogeneously distributed and scarce (Paxton *et al*., 2012; Lubbe *et al*., 2015). The outcome of this is that the true Area of Occupancy (AOO) for the Clanwilliam sandfish is confined to a 19 km stretch of river in the Oorlogskloof Nature Reserve with effective area of 0.19 km^2 (Lubbe *et al*., 2015). The Oorlogskloof Nature Reserve sandfish subpopulation is critical in terms of the survival of this highly threatened species, as it is the only viably recruiting subpopulation remaining. This makes the species as a whole extremely vulnerable to extinction (Lubbe *et al*., 2015).

Currently conservation initiatives only protect and manage populations of the two larger Cyprinids in the catchment (the Clanwilliam yellowfish and sawfin). The conservation initiatives for Clanwilliam yellowfish and Clanwilliam sawfin are however not transferable to Clanwilliam sandfish and will not secure populations of Clanwilliam sandfish. Clanwilliam thus does not have a conservation plan able to secure the safety of future populations, despite being ranked as one of the most threatened species and considered a high conservation priority by CapeNature (Impson and Swartz, 2007).

A survey in 2010 by CapeNature and the Northern Cape Department of Environment and Nature Conservation (DENC) confirmed the introduction banded tilapia, *Tilapia sparrmanii* into the municipal dam in Nieuwoudtville and have subsequently invaded Clanwilliam sandfish breeding habitat in the Oorlogskloof River. It is of great concern that were bass or bluegill sunfish introduced in a similar manner it will render this most crucial reproductive habitat unfit for Clanwilliam sandfish (Paxton *et al*., 2012; Lubbe *et al*., 2015).

No conservation measures have been directed specifically towards conserving the Clanwilliam sandfish in the past. However, the only known viable breeding population occurs in the OKNR where surveys have been conducted by the reserve staff since 2000 (Paxton *et al*., 2012). A coordinated set of actions is required that targets landowners, governing authorities including DWA, Water User Associations (WUAs), organised agriculture and angling bodies to promote sustainable land and water use practices in the catchment and to control the spread of invasive aquatic species. An active annual monitoring programme of the river has been initiated in 2010. In order to formalise conservation actions for this species in the rest of its distribution range, a Biodiversity Management Plan for Sandfish was drafted in 2012, which identified a list of potential conservations actions, along with potential implementing agents and timelines (Paxton *et al*., 2012).

1.4.2 Goals of the BMP-S:

In order to ensure the future survival of the species in the wild, further study needs to be done regarding the biology and ecology of *L. seeberi* in order to quantify the impact of habitat loss, fragmentation and predation by alien fishes on its survival. Establishment and maintenance of refuge populations in alien-free areas. Establishing a conservancy on the Oorlogskloof River, linking private land and the Oorlogskloof Nature Reserve. To achieve this, the following conservation efforts were recommended by Paxton *et al*., 2012:

(i) Elevating its status as a flagship species of the Doring River – one of the last major free-flowing rivers in the country:

(ii) Consolidate extant populations by reducing the risks of further invasions by alien fish species, especially in the Oorlogskloof-Koebee Management Unit;

(iii) Reducing the risks posed by increasing water demand and unsustainable land management practices in all catchments that fall within its distribution range;

(iv) Increasing knowledge of its biology and ecology and applying this knowledge to adaptive management strategies.

1.4.3 Benefits of the BMP-S:

The Clanwilliam Biodiversity Management Plan will set forth guidelines for the effective management of *Labeo seeberi*, reducing the likelihood of future alien fish invasions to secure future generations of sandfish. The actions proposed in this document is then set to benefit endemic fish assemblages by broadening its objectives to other affected fish assemblages downstream (Paxton *et al*., 2012).

1.4.4 Anticipated Outcomes:

The BMP-S wishes to achieve a greater awareness among landowners of the threat that sandfish is under and be more informed on how to implement ecologically sustainable land

and water use practises. Also, to draft and implement an alien fish management plan to minimise further introductions to the Olifants-Doring River system, to reclaim priority habitats and to monitor and halter the distribution of these alien fish. Lastly, the increase of knowledge regarding sandfish biology and ecology for implementation in effective adaptive management strategies (Paxton et al., 2012).

1.5 Conservation genetics:

Freshwater fish are increasingly being threatened by habitat destruction, invasion of non-native species and global climate change. This has resulted in the global decline of freshwater fish biodiversity (Leidy and Moyle, 1998; Pauly and Zeller, 2016). One of the main challenges for successful conservation strategies is to identify species/populations that are able to adapt to these environmental changes and those species/populations that will require intervention (Martinez *et al*., 2018). The ability of a population to adapt is determined by the genetic makeup of the individuals within that population. Therefore, genetic considerations are used to design effective conservation programs that ensure the survival of the species and avoid artificial selection and inbreeding depression (Vrijenhoek, 1998). Genetic markers such as microsatellite loci and mitochondrial DNA (mtDNA) have successfully been used in a number of studies regarding the conservation genetics of freshwater fish (Vrijenhoek, 1998; Abdul-Muneer, 2014; Scribner *et al*., 2016). Using these markers, researchers could identify genetic diversity within populations, genetic structure between populations and gene flow between populations. These genetic markers also help to resolve difficult taxonomic problems and delineation of possible sub-species (Vrijenhoek, 1998; Ramoejane, 2016).

Genetic diversity is a metric that measures within population variability in alternate alleles (Hughes *et al*., 2008). The ability to genetically respond and adapt may be related to both heterozygosity and the number of alleles within the population (Allendorf, 1986; Frankham, Bradshaw and Brook, 2014). In contrast, decreased population viability and increased extinction likelihood, especially in populations residing under stressful environmental conditions, can be the result of reduced genetic diversity (Reed and Frankham, 2003;

Vanderwoestijne, Schtickzelle and Baguette, 2008; Markert *et al*., 2010). Understanding how patterns of genetic diversity vary across populations could help inform predictions regarding which populations are likely to adapt in response to future disturbance while simultaneously identifying populations that might be susceptible to extinctions (Reed and Frankham, 2003; Stockwell, Hendry and Kinnison, 2003).

Another metric used for determining the conservation status of a species/population is increases and decreases in census population size as used by the IUCN (IUCN, 2018). Reduction in census population size is often associated with a decrease in genetic diversity (Willoughby *et al*., 2015). This is especially true for threatened or endangered species as compared to non-threatened taxa (Spielman *et al*., 2004; Willoughby *et al*., 2015). A population's genetic adaptability to a changing environment is thus a function of the variance of genes (allelic diversity) in a population and the number of individuals in a population, the aforementioned are thus functions of effective population size (N_e) (Ellstrand and Elam, 1993; Hare *et al*., 2011). Small, isolated populations often have very low effective population sizes (Hare *et al*., 2011). Two genetic consequences of having small population size are pronounced effects of genetic drift and increased inbreeding (Thomaz, Christie and Knowles, 2016).

Genetic drift and inbreeding lead to an increase in homozygosity. Therefore, generally small N_e leads to increased homozygosity (decreased genetic diversity) thereby ultimately reducing the adaptive potential of the population (Vrijenhoek, 1998). Another concern of increased homozygosity is the decrease of fitness of individuals within the population as result of inbreeding depression (Charlesworth and Willis, 2009). Inbreeding depression is caused by either increased homozygosity for partially recessive detrimental mutations, or increased homozygosity for alleles at loci with heterozygote advantage/ overdominance (Vrijenhoek, 1998; Charlesworth and Willis, 2009).

Population size (N) and more importantly effective population size (N_e) thus have a significant effect on allele frequencies and the rate at which they change across successive generations (Ellstrand and Elam, 1993). The larger a population, the more stable the allele frequencies are over time and by implication the more diversity can be maintained. Therefore, small, endangered populations are at an increased risk of extinction (Nei *et al*., 1975; Frankham, 2003; Frankham, 2005).

However, when populations are small, but not isolated, gene flow (in the absence of strong selection) works to counteract the increase in homozygosity caused by random genetic drift and inbreeding, thereby maintaining or regaining genetic diversity and adaptability (Wright, 1931; Palstra and Ruzzante, 2008; Martinez *et al*., 2018). This is a result of gene flow bringing in new variation, restoring alleles that were lost due to of genetic drift. The restored heterozygosity can also lead to heterosis on the phenotypic level (counteracting inbreeding). Gene flow can thus resurrect populations undergoing inbreeding depression by outcrossing with other populations of the same species. Using populations that are genetically more similar are even more effective at restoring populations as they have similar ecology profiles, thereby increasing the chances of restoration. (Westemeier *et al*., 1998; Vila *et al*., 2003; Frankham, 2015). This emphasises the importance of understanding the population structure of *Labeo seeberi*.

So the most likely scenario for the *Labeo seeberi* populations, which are known to be fragmented and small in number is as follows - Reduction in contemporary gene flow due to ongoing habitat fragmentation will likely increase the prevalence of genetic stochasticity, which in turn will negatively impact the overall genetic health and adaptability of the population (Palstra and Ruzzante, 2008; Ostergaard *et al*., 2003; Consaegra *et al*., 2005; Fraser *et al*., 2007; Schmeller and Merila, 2007; Watts *et al*., 2007). It is also known that genetic drift increases in effect, the smaller the population or N_e is. There is however a negative log-linear correlation between gene flow and population size indicating that migration or gene flow may indeed be higher into small populations, thereby counteracting the increased effect of genetic drift in maintaining genetic diversity and population viability (Ostergaard *et al*., 2003; Consaegra *et al*., 2005; Fraser *et al*., 2007; Schmeller and Merila, 2007; Watts *et al*., 2007; Palstra and Ruzzante, 2008).

Practical examples of genetic markers used in fish conservation studies include the study of Ramoejane, 2016 in which he used three genetic markers namely cytochrome oxidase 1 (*COI*) (mitochondrial), cytochrome b (*Cytb*)(mitochondrial) and recombination activating gene 1 (*Rag1*)(nuclear) to determine the evolutionary relationships of African *Labeo* spp. by clarifying the phylogeny of these species. The study also identified isolated lineages and further sub-lineages of these isolated lineages in *Labeo umbratus* using the mitochondrial gene *Cytb* in

conjunction with the nuclear S7 ribosomal protein gene intron 1 (*S7*). The study concluded that *Labeo umbratus* should be managed as several separate evolutionary significant units (ESU) and management units (MU) and not as a singular panmictic population. Studies were also done on Indian/Bangladesh *Labeo* species namely *Labeo calbasu* and *Labeo rohita*. These studies made use of microsatellite loci to determine genetic diversity within populations and differentiation between populations. This was done in order to assess the genetic differences among wild populations and the effect of aquaculture on wild populations in order to effectively conserve these resources (Singh *et al*., 2012; Hasan *et al*., 2013; Sahoo *et al*., 2014). The last example study is of the Clanwilliam rock catfish, *Austroglanis gilli*, *Austroglanis barnardi and Barbus erubescens,* fish species that co-occur with *Labeo seeberi*. The researchers used mitochondrial genes to access genetic variation within and among populations. From this data, they then identified populations that were most valuable to conserve and made recommendations for priority actions for genetic management of these species (Swartz, 2013).

It is thus important to gather as much information about the fish biology, distribution, genetic diversity and genetic structure between populations. This information will then aid conservation planners in the proper management of these fish and ensure their survivability for generations to come. The failure to produce a proper management plan may result in the loss of the Clanwilliam sandfish, and further decrease global biodiversity.

1.6 Aims and Objectives of the study:

This study aimed to genetically characterise *Labeo seeberi* in terms of historical context, and contemporary population dynamics and viability. To achieve this aim, the following objectives were set:

- Identifying species relatedness of the genus *Labeo*, based on available data. Then using the 'superficial genetic relationships' to infer possible biological traits (Chapter 2).
- Assess 'historical' and contemporary population dynamics, and population genetic diversity within the Oorlogskloof Nature Reserve (OKNR), Rietkuil (Riet) and Bos populations, using mtDNA- and nuclear microsatellite markers (Chapter 3).

Chapter 2: Species Relatedness within the Genus *Labeo* (Order: Cypriniformes, Family: Cyprinidae) using mitochondrial markers *CO1* and *Cytb* in order to infer possible biological traits

2.1 Abstract

The genus *Labeo* is a large group of freshwater fish distributed across southern Asia, the Middle East and Africa. The genus is characterised by their specialised mouth and lip modifications needed for benthic feeding. This study aimed to infer possible biological traits of *Labeo seeberi*, using the 'superficial genetic relationships' among species, based on the species relatedness within the genus *Labeo.* Nucleotide sequences for Cytochrome oxidase subunit 1 (*CO1*) (478bp) were obtained from 34 *Labeo* species and Cytochrome b (*Cytb*) (275bp) from 24 *Labeo* species were obtained. Neighbor-joining and Maximum Likelihood trees were constructed to show phylogenetic relationships. For *Cytb* the results did not follow any pattern and for the purposes of this study were uninformative. For *CO1* the tree showed two major clades namely an African clade and an Asian clade. The African clade also recovered the LNG, LFG and LCG groups proposed by Reid, 1985. According to this study *L. seeberi* is most closely related to *L. vulgaris*, but extrapolating the data using the groups proposed by Reid, 1985, *L. seeberi* and *L. capensis* are the closest living relatives for inferring biological history.

2.2 Introduction:

The need for a conservation programme for *Labeo seeberi* is urgent, as no other conservation initiatives for other Cyprinid species in the same catchment, such as the Clanwilliam yellowfish and Clanwilliam sawfin, are able to successfully secure populations of sandfish as sandfish populations are more fractured and smaller than the other two species and distributions do not completely overlap (Impson and Swartz, 2007; Paxton *et al*., 2012). Unfortunately, due to its scarcity and conservation status, sandfish are severely under-studied, and little is known about its biological traits. Therefore, many traits needed for

establishing a conservation management plan are unknown, with little to no literature available. This study thus attempted to gain more information using common ancestry in determining traits, such as time to maturity and growth rate. This will then be added to the already known biological information, such as spawning season, habitat, migration and size at maturity. This information will then aid in the drafting of a sustainable Biological Management Plan for Species (BMP-S) for sandfish (Paxton *et al.*, 2012)**.** The BMP-S aims to secure sandfish populations by reducing the threat of alien invasive fish in critical areas of historical distribution and, to more importantly, increase the knowledge of sandfish biology and ecology in order to apply this knowledge to adaptive management strategies (Paxton *et al*., 2012; Lubbe *et al*., 2015).

The genus *Labeo* is widely distributed throughout freshwater rivers and streams of Africa and Asia (Figure 2.1) (Yang *et al.*, 2012; Zheng *et al.*, 2012). Previous phylogenetic studies found that species from Africa and species from Asia clustered separately when constructing a phylogenetic tree with a clear divide between the two regions (Yang *et al*., 2009; Lowenstein *et al*., 2011; Yang *et al*., 2012). Further study using biogeography suggest that *Labeo* spp. originated in south-east Asia and then dispersed to east Asia, Africa and south Asia (Yang *et al*., 2009; Zheng *et al*., 2012). *Labeo* spp. thus spread from south-eastern Asia (Indo-China) westward through India, then Arabia and into Africa. *Labeo* spp. entered Africa through a single colonisation event, where it proceeded to spread south and eventually down to South Africa (Yang *et al*., 2009; Zheng *et al*., 2012).

Figure 2.1: Distribution of *Labeo* spp. across Africa, India and southeast Asia. Yellow arrows indicate the migration of the genus from eastern Asia to southern Africa

Currently there are nine *Lebeo* species inhabiting South African Waters, however only two of these reside in the Cape Floristic Region (CFR), namely *Labeo seeberi* and *Labeo umbratus* (Paxton *et al*., 2012; van Rensburg, 1998). Furthermore, the general paucity of genetic information for South African species means that the phylogenetic placement of South African species have not yet been fully investigated.

The aim of this chapter was thus to use the mitochondrial gene Cytochrome Oxidase 1 (*CO1*) and Cytochrome b (*Cytb*) as genetic markers to assess 'broader scale' species relatedness of the genus *Labeo.* In doing so identify the closest relatives to *L. seeberi* and superficial genetic relationships amongst these close relatives, shedding light on potentially shared biological characteristics that might be inferred and useful for conservation planning.

2.3 Materials and Methods:

Sample collection: Ethical clearance and permits were allocated internally at CapeNature. All capturing and sampling of fish was done by representatives of CapeNature. The samples were then delivered to this study. A total of 20 *Labeo seeberi* samples were collected from Oorlogskloof River, using a combination of seine nets, fyke nets and electric fishing. A small piece of the fin clip was collected from each specimen upon which the specimen was set free. The fin clips were immediately stored in 99.9% ethanol until DNA extraction. GPS co-ordinates were also logged for every specimen and deposited on the CapeNature Databank.

DNA Extraction: Tissue (± 3g) was ground in a 1.5ml Eppendorf tube. Extractions were performed using an adjusted protocol as described by Justesen *et al.* (2002). CTAB extraction buffer (2% (w/v) CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris-HCl, and 0.2 % (v/v) β-mercaptoethanol) was pre-heated at 65˚C and 600 µl was added to the ground tissue. In addition to the extraction buffer, 20 mg/ml Proteinase K was added to each sample tube and incubated overnight at 65˚C. One volume of chloroform was added to each tube and gently mixed by inversion. Then centrifuged at 13 200 rpm and 18 ˚C for seven minutes. The aqueous layer (the top layer) was transferred to a clean 1.5 ml Eppendorf tube. 200 µl Ice cold (-20˚C) isopropanol was added to each tube and mixed by inversion. Tubes were then incubated overnight at -20˚C. The formation of a pellet was produced by centrifuging samples at maximum speed for 15 minutes at room temperature. The DNA pellet was rinsed with 70% ethanol, air-dried and dissolved in 50 µl MilliQ water. Samples were left at room temperature for 30 minutes and subsequently stored in the freezer at -20˚C until further analysis.

Polymerase Chain Reaction: Polymerase chain reaction (PCR) was used to amplify the Cytochrome oxidase 1 (*CO1*) and Cytochrome b (*Cytb*) mitochondrial regions, using the primer set listed in Table 2.1. Each reaction had a total volume of 20µL consisting of 1X KAPA Ready Mix (KAPA Biosystems), 0.4µM of each primer, and 20ng template. Cycling was performed using a Veriti cycler (Lifetechnologies) using the follwing cycling conditions: initial denaturing at 95°C for 5min followed by 40 cycles of 95°C for 30sec, 60°C for 40sec and 72°C for 50sec

with a final extention step at 72°C for 5min. PCR amplicons and a 1kb DNA-ladder were loaded onto a 1.5% TBE agarose gel for agarose gel electrophoresis. Fragments were visualized under uv-light using EtBr to determine presence of bands, size and quality and negative controls included to ensure no contamination.

Table 2.1 List of the mitochondrial region amplified, the primer sequences to do so, the annealing temperatures at which it was done, the length of the fragment amplified and the source reference to the primer sets.

Mitochondrial region	Primer Sequence	Annealing Temperature (°C)	Size Range	Reference
Cytochrome oxidase 1 (*CO1*)	FF2d: 5'-TTC TCC ACC AAC CAC AAR GAY ATY GG-3' FR1d: 5'-CAC CTC AGG GTG TCC GAA RAA YCA RAA-3'	60	±609	Ivanova *et al*. 2007
Cytochrome b (*Cytb*)	L14841: 5'-AAA AAG CTT CCA TCC AAC ATC TCA GCA TGA TGA AA-3' H15149: 5'-AAA CTG CAG CCC CTC AGA ATG ATA TTT GTC CTC A-3'	55	±302	Kocher *et al*. 1989

Sequencing: Sequencing reactions were performed in the forward direction using BigDye® Terminator v3.1 sequencing kit (Applied Biosystems) as per manufacture's specifications and capillary electrophoresis was performed on an ABI3730xl sequencer (Applied Biosystems) at the Central Analytical Facilities (CAF) at Stellenbosch University.

Sequence Alignment and Phylogenetic analyses: Raw sequences of *L. seeberi* for both *CO1* and *Cytb* were edited in *Geneious software v7.1* (Kearse *et al*., 2012). All other *CO1* and *Cytb* sequences for *Labeo* species were then downloaded using the NCBI database function in *Geneious v7.1*. (Addendum A and Addendum B). Duplicate sequences were identified and removed from the list. A crude alignment was then made using the Multiple Sequence Comparison by Log-Expectation (MUSCLE) (Edgar, 2004) algorithm, implemented in the *Geneious software.* Sequences that shared no or very little overlap were discarded (incorrectly labelled or sequenced a different part of the gene as only partial coding sequences were used and not the full gene). Species specific alignments were then made and

exported separately. The number of haplotypes (N_h) for each species and assignment of individuals to each haplotype, were determined using the software package *DnaSP* version 5 (Librado and Rozas, 2009). A single individual of the most abundant haplotype was then selected to represent each species. The best substitution model was tested for, for both *CO1* and *Cytb* using the implemented function in *Mega v 6* (Tamura *et al*., 2013). Neighbor-Joining and Maximum Likelihood trees were then constructed, using the best fit substitution model based on the Bayesian Information Criterion (BIC), with the bootstrap method and 1000 replicates in *Mega v 7*.

2.4 Results:

After editing and reviewing sequence quality 13 sequences of *L. seeberi* were included for the remainder of the study for *CO1* and *Cytb*. Cytochrome oxidase 1 sequence length was 609bp, whilst for *Cytb* the amplified sequence was 302bp. The combined dataset for *CO1* after the addition of the NCBI database sequences were 167 sequences representing 34 *Labeo* spp. and for *Cytb* it was 158 sequences representing 24 *Labeo* spp. For the combined dataset, the *CO1* alignment contained 478 characters and *Cytb* 275 characters. After these alignments were further trimmed down to only represent a single individual of each of the species, 34 sequences for *CO1* and 24 sequences for *Cytb,* remained for use in the phylogenetic study. The best fit models for *CO1* were Tamura-3-Parameter + G (Neighbor-joining) and HKY +G +I (Maximum Likelihood), whilst for *Cytb* it was Tamura-Nei+ G (Neighbor-joining) and HKY +G (Maximum Likelihood). Of note is that bootstrap values for all four trees were low.

Neighbor-Joining and Maximum Likelihood phylogenies showed similar topologies for Cytochrome oxidase 1 and Cytochrome b respectively. However, taxa representation between *CO1* and *Cytb* differ, hence Neighbor joining and Maximum Likelihood trees from both *CO1* and *Cytb* are presented (Figure 2.2 – 2.5). For *CO1* (Figure 2.2 and 2.3) all African species formed a well supported group and were shown to be distinct from Asian *Labeo* species. As for *Cytb* (Figure 2.4 and Figure 2.5), no clear distinction between African and Asian *Labeo* species could be made.

In both Figure 2.2 and Figure 2.3 *L. seeberi* is most closely related to *Labeo vulgaris* (also known as *Labeo niloticus*, nile carp) which is a species from north eastern Africa (Egypt, Ethiopia and Sudan)(Azeroual *et al*., 2010). It is also worth noting that *Labeo vulgaris* closely resembles *Labeo horie* in morphology and is closely grouped in Figure 2.2 and 2.3. Unfortunately, not much biological data is available on *Labeo horie* and no comparisons could be made. *Labeo horie* and *Labeo senegalensis* paired together with very high posterior probabilities (P≥ 99%). This confirms the findings of Yang *et al*., 2012 in which they resolved *L. horie* and *L. senegalensis* as sister species. The grouping of *L. horie, L. senegalensis, L. altivelis* and *L. weeksii* also mirror that of the proposed *Labeo niloticus* group (LNG) by Reid, 1985. The grouping of *L. forskalii*, *L. annectens*, *L. parvus*, *L. simpsoni*, *L. nasus* and *L. quadribarbus*, follow that of the *Labeo forskalii* group (LFG)(Reid, 1985).The relationship of *L. nasus*, *L,. parvus*, *L. quadribarbus* and *L. simpsoni* is consistent with the findings of Lowenstein *et al*., 2011. *Labeo coubie* and *Labeo longipinnus* represent the *Labeo coubie* group (LCG) proposed by Reid, 1985 and have been confirmed sister species by Ramoejane, 2016. Of note is that Ramoejane described *Labeo batessi* as one of the most divergent species of *Labeo* and *Labeo vulgaris* and *Labeo ruddi* (not represented) as the lineage where the two species are most divergent from one another.

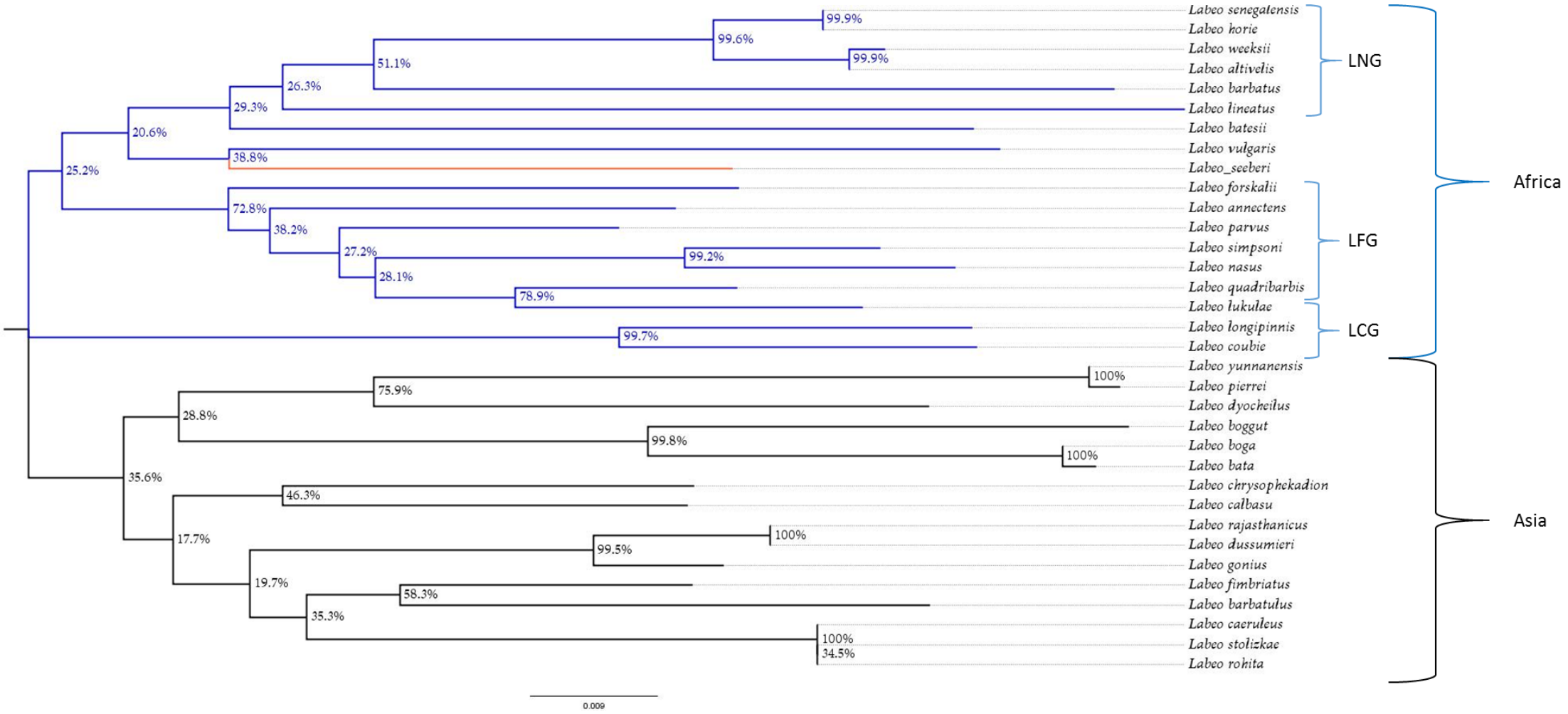


Figure 2.2: Neighbor-Joining tree depicting the phylogenetic relationship between *Labeo* spp. based on the *CO1* mitochondrial gene. The tree was constructed using the Tamura-3-Parameter + G with 1000 bootstrap repetitions. Bootstrap values are indicated on each node as percentages. Two clades are visible, the African clade in blue, and the Asian clade in black. *Labeo seeberi* is most closely related to *Labeo vulgaris* another African species, based on *CO1*. The LNG, LFG and LCG as proposed by Reid 1985 is also indicated.

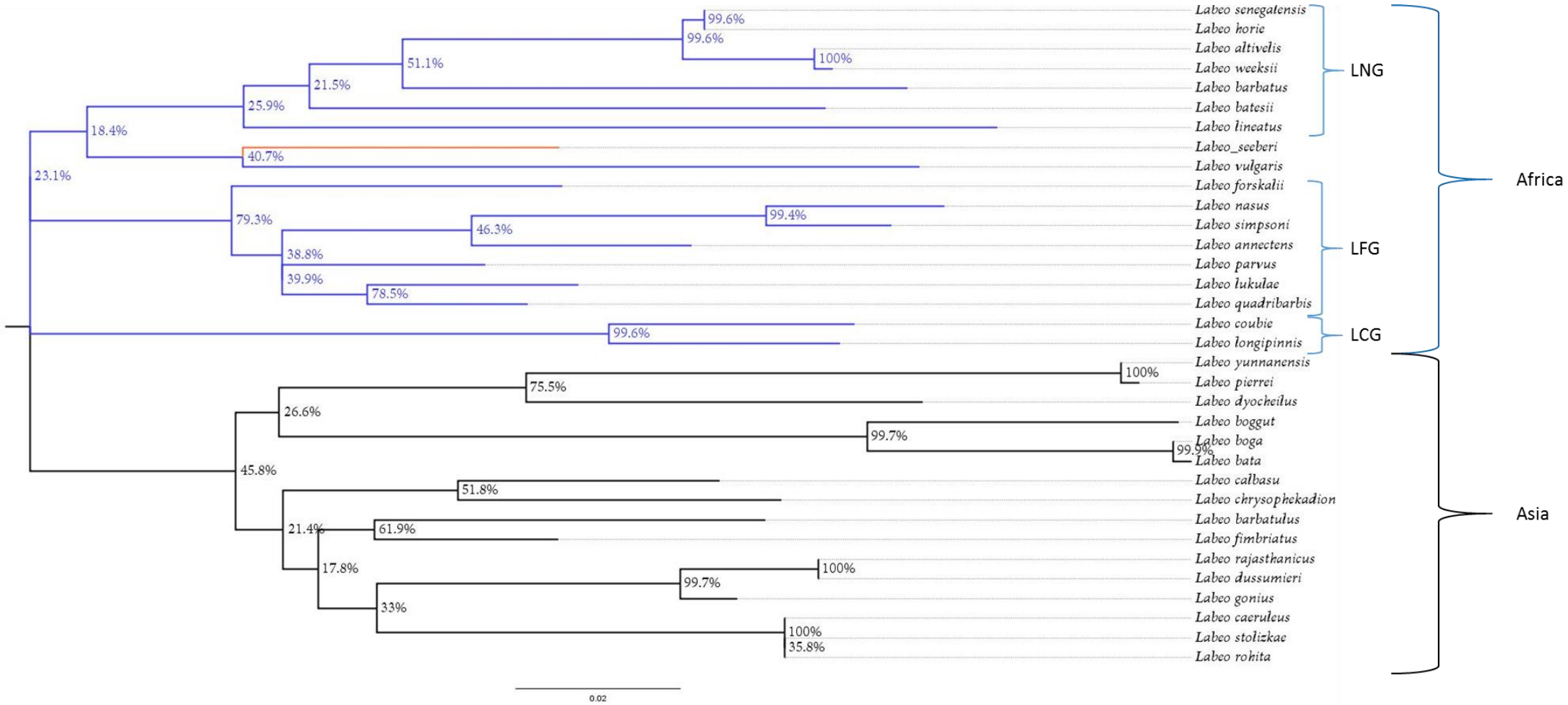


Figure 2.3. Maximimum Likelihood tree depicting the evolutionary relationship between *Labeo* spp. Based on the *CO1* mitochondrial gene. The tree was constructed using the HKY +G +I model with 1000 bootstrap repetitions. Bootstrap values are indicated on each node as percentages. *CO1* shows clear distinction between Africa (Blue) and Asia (Black). This figure shows *Labeo seeberi* to be most related to *Labeo vulgaris*, of African descent. The LNG, LFG and LCG as proposed by Reid, 1985 is also indicated.

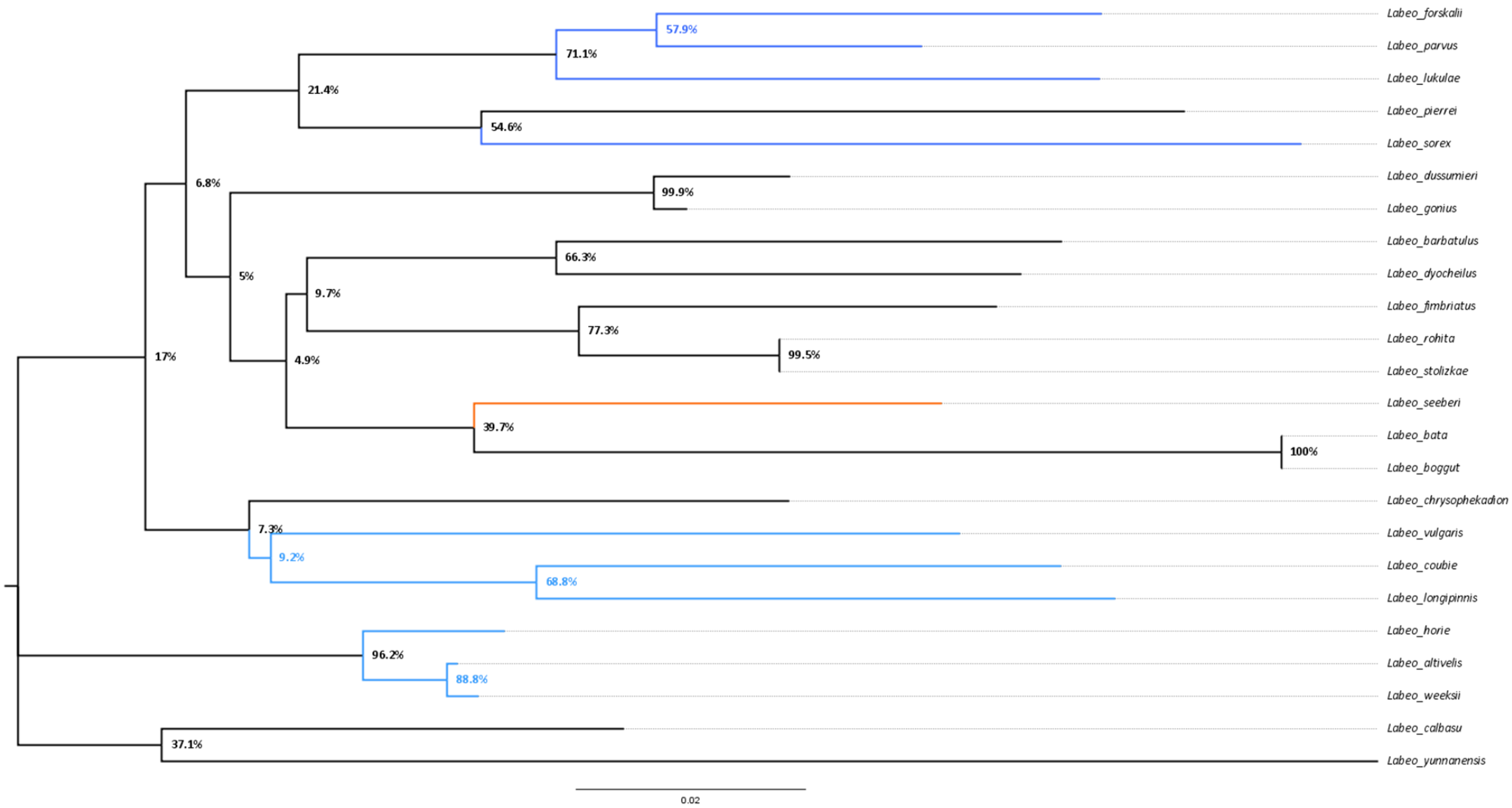


Figure 2.4. Neighbor-Joining tree depicting the phylogenetic relationship between *Labeo* spp. based on the *Cytb* mitochondrial gene. The tree was constructed using the Tamura-Nei+ G model with 1000 bootstrap repetitions. Bootstrap values are indicated on each node as percentages. *Cytb* shows no clear distinction between African species (Blue) and Asian species (Black) as the diferent specie s are interspersed amongst clades. *Labeo seeberi* groups with the Asian samples based on *Cytb* and is most closely related to *Labeo bata* and *Labeo Boggut*.

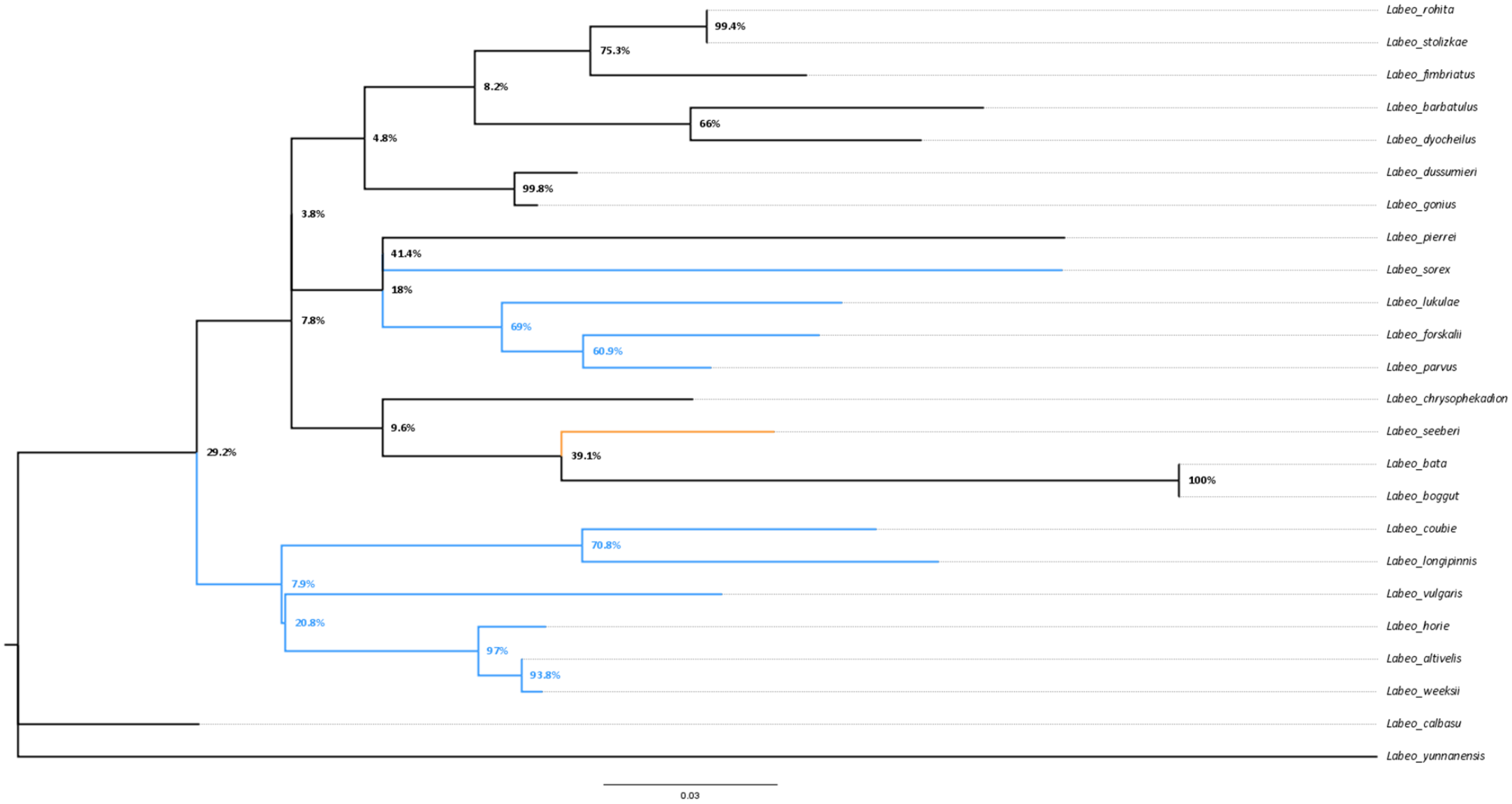


Figure 2.5 Maximimum Likelihood tree depicting the evolutionary relationship between *Labeo* spp. Based on the *Cytb* mitochondrial gene. The tree was constructed using the HKY +G model with 1000 bootstrap repetitions. Bootstrap values are indicated on each node as percentages. *Cytb* shows no clear distinction between Africa (Blue) and Asia (Black) with individuals from both continents shared in both major groupings. This figure shows *Labeo seeberi* most related to *Labeo bata* and *L. Boggut*, both of Asian descent

2.5 Discussion:

The Bootstrap values being low for all four trees indicate that the branches are not very well supported. Results must thus be interpreted with caution and cannot be seen as definitive rather as rough estimates. The relatedness analysis was in accordance with the proposed lineages of *Labeo niloticus* group (LNG), *Labeo forskalii* group (LFG) and *Labeo coubie* group (LCG) by Reid, 1985. It also supports the provisional finding of Lowenstein *et al*., 2011 in that African *Labeo* were monophyletic except for *Cytb* data, which groups the Asian *Labeo* with *Labeo seeberi*. The failure of *Cytb* might be as result of mtDNA saturation as other studies have confirmed the Asian lineage to be distinct from the African lineage (Lowenstein *et al.*, 2011; Yang *et al*., 2012).

The clear African and Asian clade represented in both Figure 2.2 and Figure 2.3 corroborates the findings of Yang *et al*., 2009; Lowenstein *et al*., 2011; Yang *et al*., 2012 (Addendum C). The above mentioned literature indicate the origin of the *Labeo* spp. to be in South East Asia, which then spread west to India, Arabia and then to Africa. Having spread to Africa in a single colonisation event, the migration continued to the south of Africa. Both Figure 2.2 and Figure 2.3 indicated *Labeo vulgaris* as the closest relative. From Figure 2.2 and 2.3 it is also apparent that *Labeo vulgaris* forms part of the LNG with *Labeo horie*, *Labeo senegalensis* and others. *Labeo vulgaris* (*Labeo niloticus*) is a freshwater fish inhabiting north-eastern Africa namely Egypt, Ethiopia and Sudan. Currently it is listed as least concern in terms of conservation priority as it has a widespread distribution with no major threats (Azeroual *et al*., 2010). It is predominantly an herbivorous species, feeding of diatoms, blue green algae and lo a lesser extent organic debris. Sexual maturity is around two years old (220-260mm TL) and spawning takes place in running waters during May to June (spring in the northern hemisphere). *Labeo vulgaris* grows to ±47 cm max length.

Using the basic information on the species together with anecdotes and *Labeo seeberi* is rheophylic, adapted to fast flowing headwaters or tributaries of large river system, undergo migrations from feeding grounds to spawning grounds during September to November (spring in the southern hemisphere) are benthic feeders and reach sexual maturity at 250mm in length etc. (van Rensburg, 1998; Paxton *et al*., 2012).

914 Extrapolating from Ramoejane, 2016 results, using the data from Lowenstein *et al*., 2011 and
915 Yang *et al*., 2012, *Labeo seeberi* belongs to the *Labeo umbratus* Group (LUG) with *Labeo*
916 *umbratus*, *Labeo capensis* and *Labeo rubromaculatus*. These fish are endemic to southern
917 Africa and members more closely resemble Asian species in terms of their morphology and
918 small scale size, than they do African Species (Reid, 1985). *Labeo seeberi* possesses the highest
919 number and smallest scale size of any African *Labeo* species (Reid, 1985). Ramoejane, 2016,
920 finally resolved *L. seeberi* to be sister species to *L. umbratus* and *L. capensis*. We can now
921 speculatively infer approximate growth rate and time to maturity of *Labeo seeberi* using one
922 of its South African relatives, which not only is geographically close by, but also shares a
923 similar habitat namely *Labeo capensis*. Like *Labeo seeberi*, *Labeo capensis* also inhabits deep
924 silted pools mixed with fast flowing rocky rapids. Like the CFR the Caledon River in which
925 *Labeo capensis* makes its home also has decreased water-flow during the dry season leading
926 to crowded isolated pools. *Labeo capensis* also spawns during the same time as *L. seeberi* and
927 reaches sexual maturity at the same size (250mm) and breeds in summer over shallow rocky
928 rapids, where they aggregate in large numbers (Skelton, 2001). Using this as reference, we
929 can speculatively infer that *Labeo seeberi* reaches sexual maturity at approximately ±4 years
930 of age growing 4-6 cm per year up unto the sixth year after which it decreases steadily.
931 Females grow faster than males, and have a longer lifespan (Baird, 1976). These are however
932 only preliminary, speculative results and further testing needs to be done to provide definitive
933 answers.

934

935 2.6 Conclusion:

936

937 Using *CO1* as marker correctly identified three of the five African *Labeo* groups proposed by
938 Reid, 1985. Identified that *Cytb* is not a suitable marker for determining *Labeo* phylogeny as
939 it has potentially undergone mtDNA saturation. By determining the closest common ancestor
940 to *Labeo seeberi* we have added to the preliminary biological knowledge of the species and
941 thereby aid the BMP-S that is being drafted for sandfish conservation. This includes the
942 inferred approximate growth rate of *Labeo seeberi* as being 40-60 mm per year and reaching
943 sexual maturity at ± 4 years old. Information that was previously unknown. We also speculate

944 that *Labeo seeberi* is an r-type spawner (mass spawning events) with short hatch times for
945 the eggs ±2 days. Knowing the growth rate is a benefit to conservation as rangers can
946 determine the time to outgrow the prey size class and therefore the period of most mortality.

Chapter 3: Assessing population genetics of the Clanwilliam sandfish, *Labeo seeberi* within its last remaining habitat, using microsatellite and mtDNA markers

3.1 Abstract

The Clanwilliam sandfish (*Labeo seeberi*) was once widespread throughout the Olifants-Doring River system, but suffered significant population declines. This has led to the absence of sandfish in the Olifants River and few populations in the Doring River. Sandfish play an important role in maintaining freshwater river habitat by controlling algae levels and cycling nutrients. Thus, sandfish need to be conserved in order to maintain river health. Gaining knowledge about the genetic diversity and population structure of *L. seeberi* is important in drafting an effective conservation and management programme to secure sustainable sandfish populations. Genetic diversity and population differentiation were estimated using 6 microsatellite loci and the hyper variable region (*D-loop*) of the mitochondrial DNA (mtDNA). For the microsatellite data, the number of alleles (A_N) ranged from 6 to 11, expected heterozygosity (H_{Enb}) ranged from 0.696 to 0.700 and the inbreeding coefficient (F_{IS}) was non-significant. There was no significant differentiation between populations (OKNR, Riet and Bos) according to the Analysis of Molecular Variance (AMOVA), Factorial Correspondence Analysis (FCA) and the F_{ST} –values. Effective population size (N_e) of Riet (118) and Bos (15) were relatively small compared to the OKNR population (465). Microsatellite data also indicate no recent genetic bottleneck and no significant relatedness (*r*) among individuals of the same population. For the mitochondrial data, haplotype diversity (*h*) was high 0.782 to 0.821 and no significant population differentiation (Φ_{ST}) detected. These results indicate that OKNR, Bos and Riet are not genetically differentiated, with some level of gene flow maintaining genetic diversity. It also indicates that OKNR is the primary breeding grounds, but that diversity outside of the OKNR exists. Considering the critically endangered status of *Labeo seeberi*, these results could be considered for future management and conservation programmes.

3.2 Introduction:

The value of conserving and researching the Clanwilliam sandfish is inestimable, not as economic commodity, but its integral function in the freshwater biodiversity of the Cape Floristic Region (CFR) (Paxton *et al.*, 2012). This integral function is the cycling of nutrients and controlling algae in the rivers they inhabit, as result of their algivorous and detritovorous feeding habits (Skelton, 2001). If these conservation practices learned from *Labeo seeberi* prove to be effective, these practices could set to benefit endemic fish assemblages by broadening its objective to other effected fish assemblages downstream (Paxton *et al.*, 2012).

Once widespread throughout the Olifants-Doring River system, much of this distribution has been lost (Paxton *et al.*, 2012). In a survey done in 2011, as part of the Biodiversity Management Plan for Species (BMP-S) aimed at the Clanwilliam sandfish, the investigators presented evidence to support an 80% decline in population numbers as opposed to pristine population. With most of the decline being attributed to the reduction/seizure of populations within the main stream of the Olifants and Doring Rivers (Paxton *et al.*, 2012; Lubbe *et al.*, 2015). This reduction in population viability within the main streams was two-fold. First was the large-scale water resource infrastructure (dams and weirs) and water abstraction in the Olifants River (Paxton *et al.*, 2012). Secondly is the fish population composition of the mainstream Doring River system. Dams and weirs create barriers to the fish ability to move freely throughout the river system, affecting the dispersal of juveniles, the migration to and from feeding/breeding grounds and overwintering areas (Paxton *et al.*, 2012). The populations that do persist in the mainstream are mainly comprised of old, larger individuals (>400mm) that are beyond the prey size class (Lubbe *et al.*, 2015). Although these populations seem to be reproducing, the offspring seldom survive to adulthood as result of predation by alien invasive fish species. Leaving little to no recruitment of juveniles. Data thus indicates *L. seeberi* being completely absent from the Olifants River and the remaining populations in the Doring River main stream becoming increasingly scarce and heterogeneously distributed, and only occurring in the middle and northern reaches (Paxton *et al.*, 2012; Lubbe *et al.*, 2015).

Apart from the scarce population persisting in the mainstream Doring River, the majority of sandfish populations are confined to the upstream reaches of small isolated tributaries

guarded against alien invasive fish species by means of natural barriers. The only known successfully recruiting sandfish populations occur in the confined reaches of the Oorlogskloof Nature Reserve (OKNR) of the Oorlogskloof River, called the Koebee in its upper reaches, in the Northern and Western Cape Provinces (Paxton, 2002; Paxton *et al*., 2012; Lubbe *et al*., 2015).

In order to secure sustainable populations of sandfish, conservation efforts need to protect distinct populations of the species and preserve their evolutionary significance. This is because this diversity could be the result of locally adaptive traits that may affect the ability of a species to respond to new evolutionary challenges (Vrijenhoek, 1998). Populations that have been historically isolated and are genetically distinct from one another are designated as Evolutionary Significant Units (ESU). For a population to be considered an ESU, it must comply to two criteria: The population has to be reciprocally monophyletic for mtDNA alleles and show significant divergence of allele frequencies at nuclear loci (Moritz *et al*., 1994). If a population does not fit both criteria, but still shows divergent allele frequencies at nuclear loci, it could still be considered for conservation as a Management Unit (MU). Evolutionary significant units thus recognise the major intraspecific genetic diversity for conservation actions (historical populations/diversity), whilst MU's represent demographically independent sets of populations characterised by restricted gene flow that can be managed to retain the larger ESU (Moritz *et al*., 1994; Palsboll *et al*., 2007; Funk *et al*., 2012). Delimiting conservation units as important elements of intraspecific diversity has significant implications for genetic rescue of small populations through restoration of gene flow or by augmentation and it is becoming increasingly clear that genetic rescue needs to be considered more broadly if increased population extinction is to be averted (Love Stowell *et al*., 2017; Ralls *et al*., 2018). It was proposed that mixing MU's, but not ESU's could be considered an appropriate strategy for genetic rescue (Moritz *et al*., 1994). The distinction between ESU and MU is important as it affects the way in which genetic evidence is obtained and interpreted and to be used effectively. ESU is concerned with the historical population structure, mtDNA phylogeny and long-term conservation needs whilst MU addresses the current population structure, allele frequencies and short term management issues (Moritz 1994). The power of genetic tools to identify ESU's and MU's greatly contribute effective conservation programs, conserving historical diversity and local adaptation. This is evident in the studies of Ramoejane, 2016 and

Swartz, 2013 (Chapter 1.5 – Conservation genetics) in which they discovered isolated lineages that are to be managed as ESU's and further sub-structuring of populations within those lineages which are to be managed as MU's. These studies also identified hybrid specimens leading to even more specialised conservation actions. Thus without the use of genetic tools in conservation, a lot of diversity and local adaptation could be lost (Vrijenhoek, 1998).

It was thus proposed that the knowledge of the species biology and ecology should be greatly increased as to apply it in adaptive management strategies (Paxton *et al*., 2012). One of the key components in the drafting the adaptive management plan is determining the genetic diversity and genetic population structuring of *Labeo seeberi* using microsatellites and mitochondrial DNA. Microsatellite markers were selected for identifying contemporary genetic processes (Vrijenhoek 1998; Abdul-Muneer, 2014). Mitochondrial markers were used to identify historical genetic processes. This study did however use a section of the control region (*D-loop*) which is not under functional constraint so that more mutations could be present in order to calculate diversity more effectively (Vrijenhoek, 1998; Scribner *et al*., 2016). This study therefore aims to quantify the genetic diversity and population genetic structure of *Labeo seeberi* in the Olifants-Doring River system using microsatellite and mitochondrial DNA markers, in order to contribute to management and restocking plans ensuring the long term survival of the species.

3.3 Materials & Methods:

3.3.1 Study populations and specimens:

Ethical clearance and permits were allocated internally at CapeNature. All capturing and sampling of fish was done by representatives of CapeNature. The samples were then delivered to this study. In total 128 individuals were sampled from three different sampling sites (Oorlogskloof Nature Reserve (82), Rietkuil (36) and Bos (10) (Figure 3.1) spanning the natural distribution of *Labeo seeberi* within the Olifants-Doring River system (Table 3.1).

Individuals were captured using a combination of sein nets and large fykes. Tissue samples (fin clips) were collected from each individual and stored in a 2.5ml tube filled with 99.9% Ethanol. Samples were stored at 4°C until DNA extractions could be performed using the adjusted CTAB protocol as described by Justesen *et al.* 2002 as described in Chapter 2.

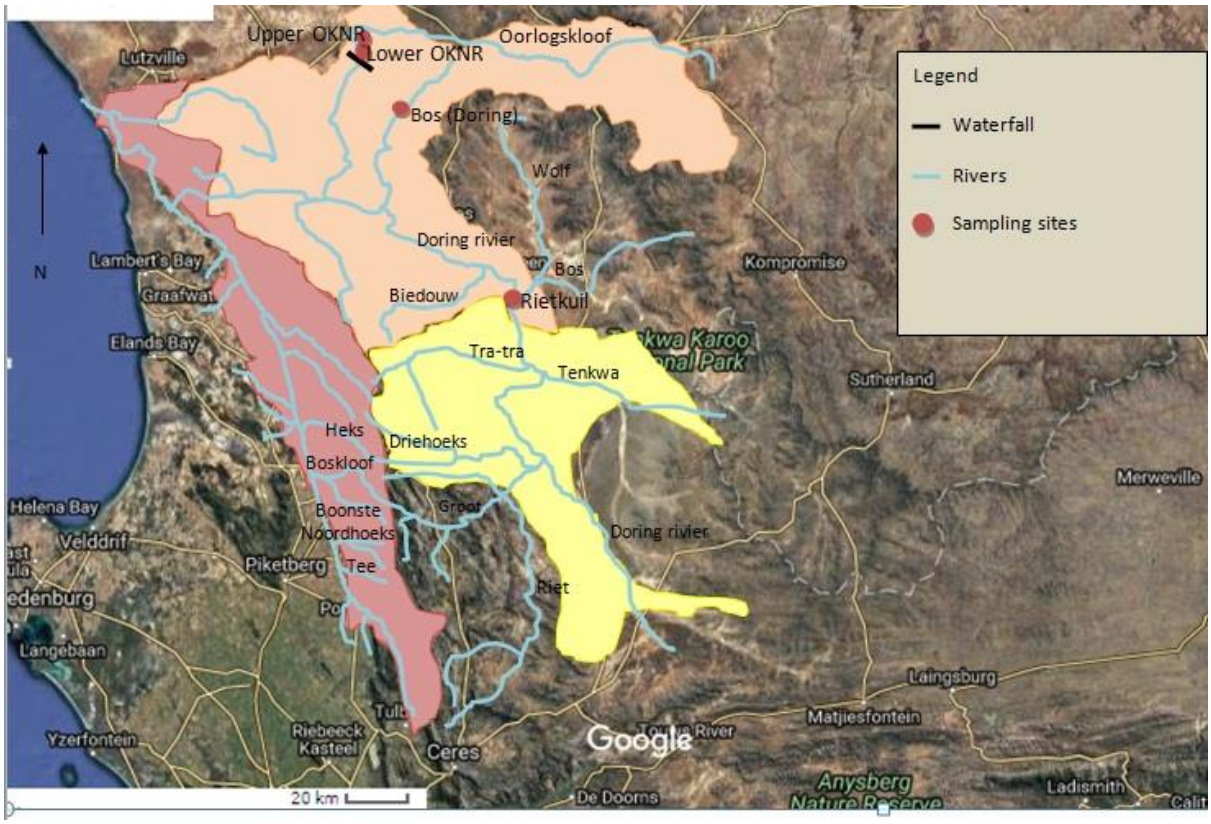


Figure 3.1: Map of the Olifants-Doring River system in the Northern and Western Cape of South Africa. The Clanwilliam sandfish is restricted to the northern reaches of the Doring River (orange range) with recruitment restricted to the Oorlogskloof-Koebee tributary River. The red range indicates where sandfish were historically present but are now extinct. The yellow range indicates where sandfish possibly still are but in low frequency with effectively zero recruitment possible. The red dots (OKNR), (Riet) and (Bos) indicate the position of the three populations.

Table 3.1: Names, symbols, co-ordinates and number of individuals for each sampling population of *Labeo seeberi* used in this study

Sampling Site		Co-ordinates	Symbol	Number of Individuals
Oorlogskloof Nature Reserve	Site 1	31°31'27.4"S 19°06'38.7"E to 31°27'38.0"S 19°04'54.0"E	OKNR Site 1	8
	Site 2		OKNR Site 2	2
	Site 3		OKNR Site 3	7
	Site 4		OKNR Site 4	8
	Site 5		OKNR Site 5	16
	Site 6		OKNR Site 6	10
	Site 7		OKNR Site 7	7
	Site 8		OKNR Site 8	8
	Site 9		OKNR Site 9	2
	Site 10		OKNR Site 10	2
	Site 11		OKNR Site 11	12
Rietkuil		31º35'01,93''S 19º04'20,04''E	Riet	36
Bos		32º07'20,90''S 19º29'13,38''E	Bos	10

Co-ordinates represented on a physical map in (Figure 3.1).

3.3.2 Microsatellite genotyping:

Eleven microsatellite markers were sourced from literature on sister *Labeo* species, *Labeo rohita* and *Labeo fimbriatus,* as indicated in Table 3.2 and Addendum D. These markers were then cross- amplified in *Labeo seeberi*. The markers that successfully amplified, were polymorphic, and consistently scoreable were selected for dowstream analysis. Genotyping of the individuals from the three different sampling regions were performed using 6 microsatellite loci Indicated with an * in Table 3.2

Table 3.2: Microsatellite loci information stating the name of locus, the fluorescent label, primer sequence, annealing temperature, the size range and the repeat sequence.

Microsatellite	Fluorecent Label	Primer Sequence	Annealing Temperature (°C)	Reference
Lro_26*	PET	F: 5'-AGA TCA TTG CTG GGG AGT GTT TAT-3' R: 5'-GAC CTG CCT GTG CCA TCT GTA-3'	58	Swain *et al.*, 2013
Lr_28	PET	F: 5'-TTC ACG GAC AGA TTT GAC CCA G-3' R: 5'-AGT CTT TTC AGG AGA TTA GCA G-3'	60	Patel *et al.*, 2009
Lr_29	NED	F: 5'-ACG TAA AGG TCA CAA GCT GAA G-3' R: 5'-AGC ACG GTG TTT GTG TGC GAG-3'	60	Patel *et al.*, 2009
Lr_30	VIC	F: 5'-ACG CGC TAG GGT CGT ACA GTG-3' R: 5'-CAG CAT CAT GTT AAG CGC TGT C-3'	60	Swain *et al.*, 2013
Lr_36*	NED	F: 5'-AGC GTG TCT GAT GTG TGA AAG G-3' R: 5'-TCA GAT GCC TCC TGC ATT CTG-3'	58	Swain *et al.*, 2013
Lr_41*	VIC	F 5'-TCC AGT CAC CAC ATG CGT TTG-3' R 5'-GTC GAT TTC ATC GTG AGG CTC-3'	58	Patel *et al.*, 2011
Lr_44	NED	F: 5'-CAC CCA GGG AGT TAG TTT CTG-3' R: 5'-AAA GAG CAT CAT GGC ATT GAC-3'	57	Patel *et al.*, 2011
Lr_46	FAM	F: 5'-TGA CGT ATT GTC AAC TAT GGT G-3' R: 5'-TCC ACC TTC AAt ACC ATG ACT G-3'	58	Patel *et al.*, 2009
LF_8*	FAM	F: 5'-GTG AAG CAA CGA CTT CAG AGA G-3' R: 5'- CCA GAA GAC CAT AGC AAC CAC-3'	53	Swain *et al.*, 2012
LF_15*	VIC	F: 5'-ACA CTC ACA CTC GCT CAC TCA C-3' R: 5'-CGG TGA ATG CTG ATG AAC TG-3'	55	Swain *et al.*, 2012
LF_16*	FAM	F: 5'-AAC GTC ACA CAT GCT CCT AGT C-3' R: 5'-CTG CCC ATG ACA CTG AAA CTC-3'	60	Swain *et al.*, 2012

All polymerase chain reactions (PCR) were performed using the Veriti Thermo Cycle (*LifeTechnologies*), with KAPA Taq Ready Mix (*KAPA Biosystems*). The PCR conditions were as follows: A final reaction volume of 10µl consisted of 1X KAPA Taq Ready Mix, 100ng of DNA and 0.4µM of each primer. The cycling conditions were as follows: Denaturing at 95°C for 5min followed by 30 cycles of 95°C for 45sec, X°C (Table 3.2) for 1min and 72°C for 2min with a final extention step at 72°C for 7min.

Capillary elctrophoresis was performed using a 3730XL Genetic Analyzer (*LifeTechnologies*). A LIZZ600 size standard (*LifeTechnologies*) was used for sizing. GeneMapper® version 4.0 (*LifeTechnologies*) was used for allele scoring and were manually revised to ensure corectness. File conversions were performed using MS tools (Kim and Sappington, 2013).

3.3.3 Microsatellite Data Analysis: Genetic Diversity

The basic genentic diversity estimates such as unbiased expected (H_{Enb}) and observed (H_O) heterozygosity, number of alleles (A_N) and allelic richness (A_R) were calculated as follows. The number of alleles (A_N) per population per locus were calculated using the software program Fstat version 2.9.3.2 (Goudet, 1995). The program HP-RARE (Kalinowski, 2005) was used to calculate allelic richness (A_R) and private allelic richness (A_{PR}), standardised for sample sizes. Genetix version 4.03 (Belkhir *et al*., 2000) was used for calculating the observed (H_O) and Nei's unbiased expected (H_{Enb}) heterozygosities (Nei, 1978) as well as the F_{is}-values. The departure from Hardy Weinberg Equilibrium (HWE) (exact probability test (Guo and Thomson, 1992), 500 batches, 10 000 iterations (Raymond and Rousset, 1995)) and the null allele frequencies (Fr_{null}) were both calculated per population per locus using the computer software Genepop version 4.0.7 (Raymond and Rousset, 1995).

3.3.4 Microsatellite Data Analysis: Population Differentiation:

Population differentiation analysis was split into two sections. The first section was to test for possible genetic structure between OKNR, Rietkuil and Bos. The importance of this was to determine if populations of different rivers were genetically distinct from one another or if they represented a signle homogeneous population. The second section was to determine whether there was population genetic structure between the 11 sampling sites within the OKNR (Table 3.1) as this region serves as the spawning sanctuary for *Labeo seeberi*. This area is thus of great importance, as any genetic substructuring would directly effect the overall genetic diversity of the species. To answer these quieries the following analysis were

conducted for both OKNR as separate from the rest, and OKNR as a whole against Rietkuil and Bos.

Global analysis across major regions: OKNR, Rietkuil and Bos

:

To assess genetic differentiation between populations, pairwise F_{ST}, Analysis of Molecular Variance (AMOVA), Factorial Correspondence Analysis (FCA) and Bayesian clustering analytical methods were calculated as follows. Pairwise F_{ST} were calculated and the significance levels set at 0.05 ($p \leq 0.05$) in the software program Arlequin version 3.5 (Excoffier and Lischer, 2010). To support the F_{ST}-data, an AMOVA was performed also using the software program Arlequin version 3.5. Genetix version 4.03 (Belkhir *et al*., 2000) was used to calculate and visually present the genetic differentiation between populations by means of a Factorial Correspondence Analysis plot. The Bayesian clustering software Structure version 2.3.4 (Pritchard, 2000; Falush, 2003; Falush, 2007; Hubisz *et al*., 2009) was employed to detect the most likely number of clusters/populations (*K*) and the proportion of each *K* within each idividual genome. Ten replicates of 50 000 runs and a 10 000 iteration burn-in period, assuming no prior population information, was performed for each *K*. The estimated log probability for each value of *K* was calculated assuming the admixture model with correlated allele frequencies and *K* ranging from *K*=1 (Panmictic) to *K*=5 (number of distinct sampling sites + 2). The web based program Structure harvester version 0.6.93 (Earl and von Holdt, 2012) was implemented to process the STRUCTURE result files and identify the most likely *K*-value using the method of Evanno *et al*., 2005. The results for the 10 replicates were averaged and the output visualised (bar-plot visualising the proportion of each *K* within each individual genome) using the web based program CLUMPAK (Kopelman *et al*., 2015)

Local analysis within the OKNR:

The same analyses were conducted, as described for the global analysis, with the exception of the following alterations to the parameters of the Bayesian clustering software, STUCTURE:

The number of possible populations were changed to *K*=15 in accordance with the change in number of sampling sites (distinct numbers of sites + 4). Ten replicates of 50 000 runs and a 10 000 burn-in period was performed for each *K,* assuming no prior population information. All other parameters remained the same.

3.3.5 Microsatellite Data Analysis: Effective population size (N_e) and population bottlenecks:

Contemporary effective population sizes were estimated for each of the sampling populations, using the linkage disequilibrium (random mating) and heterozygosity excess method as implemented in Ne Estimator v. 2.0.1. (minimum allele frequency Pcrit = 0.02) (Do *et al.*, 2014). Significance tests were set at upper and lower 95% confidence intervals. To test for recent bottlenecks, the Wilcoxin signed rank test (Luikart *et al*., 1998), under all three mutation models [Infinite alleles model (IAM), stepwise mutation model (SMM) and the two-phased model (TPM)] were carried out as implemented in BOTTLENECK v1.2.02 (Piry *et al*., 1999). Analysis were carried out using 10 000 replications at the 5% nominal level and a TPM composed of 70% SMM and 30% IAM and a variance of 30.

3.3.6 Microsatellite Data Analysis: Relatedness, *r*, per population:

Mean relatedness per population was calculated in GENALEX v6.5 (Peakall and Smouse, 2012), using the relatedness estimator, *r*, of Queller and Goodnight ,1989 (Significance testing by 999 bootstrap replicates).

3.3.7 mtDNA Analysis: Sequencing and Alignment

Thirteen individuals were selected at random from each of the three sampling regions (Fig 3.1), for a total of 39 individuals included in these analyses. A 736bp region of the control region (*D-Loop*) was amplified by means of Polymerase chain reaction (PCR) using the primer pair of Fish G(F): 5'-GCATGGGTCTTGTAATCCGA-3'and Fish F(R): 5'-TAGTAAGGTCGGGACCATGC-3' (Chung *et al.*, 2010), with a total reaction volume of 20µl

Kappa Taq Ready Mix (Kappa Biosystems). The thermal cycling profile of the PCR was 95°C for 3min, followed by 35 cycles of 30s at 95°C, 30s at 57°C and 1 min at 72°C, with a final extension step at 72°C for 7 min. The PCR amplicons were sent to the Stellenbosch University Analytical Facility (DNA sequencing unit) where samples were purified using the Qiagen gel clean-up system and sequenced using standard Sanger sequencing chemistry (BigDye® terminator v3.1 cycle sequencing kit, *Life Technologies*). Sequencing products were purified using Sephadex spin columns (Princeton Seperations, Adelphia, NJ) and analysed *via* capillary electrophoresis on a 3730XL Genetic Analyzer (*LifeTechnologies*). Sequences were aligned using the MUSCLE algorithm (Edgar, 2004), implemented in MEGA v6 (Tamura *et al*., 2011), manually corrected and trimmed to equal lengths.

3.3.8 mtDNA Analysis: Molecular diversity and population differentiation:

The following genetic parameters were estimated for each sampling region: Number of polymorphic sites (S), haplotype diversity (h) and nucleotide diversity (π) using DnaSP v5.10.1 software (Librado and Rozas, 2009). Analysis of Molecular Variance (AMOVA) and pairwise molecular differentiation (Φ_{ST}) were calculated using ARLEQUIN v3.5. (Excoffier, 2010). Fisher's exact test was done to evaluate the significance Φ_{ST} results. Φ_{ST} was estimated using the Tamura-Nei (as suggested by the AIC and BIC scores implemented in MEGA v6. modeltest function) model of nucleotide substitution, with α- value of the gamma distribution set to default 0. The significance of the Φ-statistics were determined using 1000 iterations. Maximum parsimony haplotype networks (Polzin and Daneshmand, 2003) were constructed by means of the median joining algorithm (Bandelt *et al*., 1999) with default parameters in Network v4.6.1.2.

3.4 Results

3.4.1 Microsatellite genotyping:

Of the 11 microsatellite markers listed in (Table3.2), six microsatellite markers amplified successfully and were polymorphic. These six markers include Lro_26, Lr_36, Lr_41, LF_8, LF_15 and LF_16. These six microsatellite markers were thus included for further analysis. A criteria was put in place allowing only individuals with four or more markers (> 50%) successfully typed to be included for further analysis. A total of 128 individuals were sucessfully genotyped for four or more microsatellite markers. Null allele frequencies (F_{Rnull}) are shown in Addendum E.

3.4.2 Microsatellite Data Analysis: Genetic Diversity:

Genetic diversity statistics including number of individuals (N_{ind}), number of alleles (A_N), allelic richness (A_R), Private allelic richness (A_{PR}), Hardy Weinberg Equilibrium (HWE), observed (H_O) and unbiased expected (H_{Enb}) heterozygosities and F_{IS} are shown in Addendum E and summarised as means per population in Figure 3.2

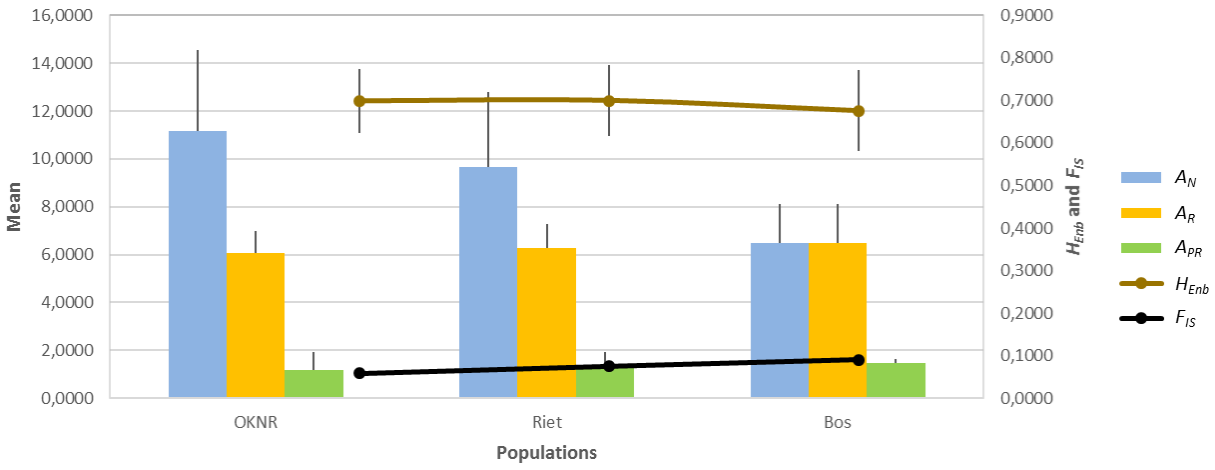


Figure 3.2: Summary of mean diversity statistics per population including number of alleles A_N, Allelic Richness A_R, Private Alleles A_{PR}, Unbiased Expected Heterozygosity H_{Enb} and The Inbreeding Coefficient F_{IS}.

Overall, the mean number of alleles per sampling region was moderate to low (Figure 3.2). Ranging from 11.167 (OKNR) as the most number of alleles (A_N) per sampling site to 6.500 (Bos) as the fewest number of alleles per sampling site. OKNR and Riet displayed similar number of alleles per sampling region whilst the number of alleles per sampling region for Bos was considerably fewer. Based on frequency and corrected for population size disparity, the allelic richness (A_R) are much more comparable between the three sampling regions, 6.069 (OKNR), 6.279 (Riet) and 6.500 (Bos). The mean unbiased expected heterozygosity for each of the three sampling regions were also comparable, 0.698 (OKNR), 0.700 (Riet) and 0.676 (Bos). Interestingly, all three sampling regions have private alleles and although the mean number of private alleles per sampling region is few, it would seem as though it has an inverse proportion to population size/sampling size. The mean Fixation indices F_{IS} for the three regions were all very low.

3.4.3 Microsatellite Data Analysis: Population Differentiation:

The F_{ST}-values for estimating genotypic differentiation between OKNR, Riet and BOS (Table 3.3 and Addendum F) ranged from 0.004 to 0.008 and suggested that OKNR is marginally differentiated from Riet (F_{ST} = 0.008; P = 0.044), whilst all other population comparisons were non-significant. Pairwise F_{ST}-values for calculating the differentiation between the 11 different sampling sites within the *Labeo seeberi* sanctuary namely OKNR (Table 3.4 and Addendum G), ranged between 0.00016 and 0.25714. Overall there were no significant differentiated sites, except for site 11 as compared to a number of the other sites.

Table 3.3 Pairwise F_{ST}- values describing the extent of genetic differentiation between 3 regions within the Olifants-Doring River system

	OKNR	Riet	Bos
Riet	0.008*	0.000	
Bos	0.004	0.005	0.000

*** Significance $P \leq 0.05$**

Table 3.4 Pairwise F_{ST}- values describing the extent of genetic differentiation between 11 sampling sites along the Oorlogskloof River.

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8	Site 9	Site 10	Site 11
Site 2	0.065	0.000									
Site 3	-0.002	0.064	0.000								
Site 4	-0.049	0.118	-0.025	0.000							
Site 5	0.003	0.097	-0.003	-0.059	0.000						
Site 6	0.000	0.061	0.033*	-0.027	-0.002	0.000					
Site 7	-0.018	0.053	-0.001	-0.045	-0.006	-0.005	0.000				
Site 8	-0.003	0.051	0.008	-0.083	0.003	-0.002	-0.007	0.000			
Site 9	0.013	0.048	-0.003	-0.010	-0.021	0.025	0.003	0.023	0.000		
Site 10	0.005	0.062	0.019	-0.002	0.001	-0.001	0.000	0.024	0.015	0.000	
Site 11	0.146*	0.257	0.085*	0.074	0.069	0.156*	0.12869*	0.132*	0.123	0.075	0.000

*** Significance $P \leq 0.05$**

The analysis of molecular variance between the three sampling regions (Bos Riet and OKNR) (Table 3.5) corroborated the pairwise F_{ST} results (Table 3.3), indicating weak overall population differentiation and finding the accompanying *F*-statistic to be non-significant (F_{ST} = 0.005, P = 0.784, 0.548% of the total variation). The majority of the total variance (90.255%) came from within the individuals, whilst the variance among individuals within populations accounted for the remaining 9.197%. The AMOVA results for the 11 sites of the OKNR (Table 3.6), mirror that of the beforementioned AMOVA as it also has weak population/ site differentiation with the accompanying *F*-statistic being non-significant (F_{ST} = 0.005, P = 1.000**,** 0.508% of the total variation). The majority of the variance was distributed within individuals (91.945%) whilst the remainder of the variance was ditributed among individuals within popupulations (7.547%). It is important to note that although the pairwise F_{ST}'s indicated significant differentiation for a few of the sampling sites (Table 3.3 and 3.4) the differentiation as a whole over all populations tested is weak and therefore not sigificant (Table 3.5 and 3.6)

Table 3.5 AMOVA results as weighted average across 6 loci of study populations (OKNR, Rietkuil and Bos) of *Labeo seeberi*

Source of variation	*Sum of squares*	*Variance components*	*% Variation*
Among populations	*6.062*	*0.011*	*0.548*
Among Individuals within populations	*285.124*	*0.194*	*9.197*
Within individuals	*242.500*	*1.900*	*90.255*
Total	*533.686*	*2.105*	
F_{IS}*: 0.092*	*P: 0.000**		
F_{ST}*: 0.005*	*P: 0.784*		
F_{IT}*: 0.097*	*P: 0.000**		

*** Statistical significance at the 95% nominal level ($P \leq 0.05$)**

Table 3.6 AMOVA results as weighted average across 6 loci of the 11 sampling sites of *Labeo seeberi* in the OKNR

Source of variation	*Sum of squares*	*Variance components*	*% Variation*
Among populations	*23.981*	*0.011*	*0.508*
Among Individuals within populations	*158.643*	*0.158*	*7.547*
Within individuals	*157.500*	*1.928*	*91.944*
Total	*340.124*	*2.097*	
F_{IS}*: 0.076*	*P: 0.004**		
F_{ST}*: 0.005*	*P: 1.000*		
F_{IT}*: 0.081*	*P: 0.001**		

*** Statistical significance at the 95% nominal level (P ≤ 0.05)**

Factorial Correspondance Analysis (FCA) (Figure 3.3) illustrates some evidence of separation of the clusters although there is still a large proportion of overlap at the edges. The blue cluster made up of individuals from the Riet population, the White cluster of individuas from Bos and the yellow cluster consisting of individuals from the OKNR. This supports both the pairwise F_{ST} (Table 3.3) and AMOVA (Table 3.5) results as the differentiation of OKNR and Riet is illustrated in the FCA-plot, but both these clusters still have a large proportion of

overlapping individuals at the edges. This means that the global population differentiation as indicated by the AMOVA is still a very small component of the variance. The FCA-plot for the 11 sites of the OKNR (Figure 3.4) appears to have no significant clusters apart from a few small outlier sites. This too supports the previosly mentioned pairwise F_{ST} (Table 3.4) and AMOVA (Table 3.6) in showing that the 11 sites within the OKNR are not differentiated.

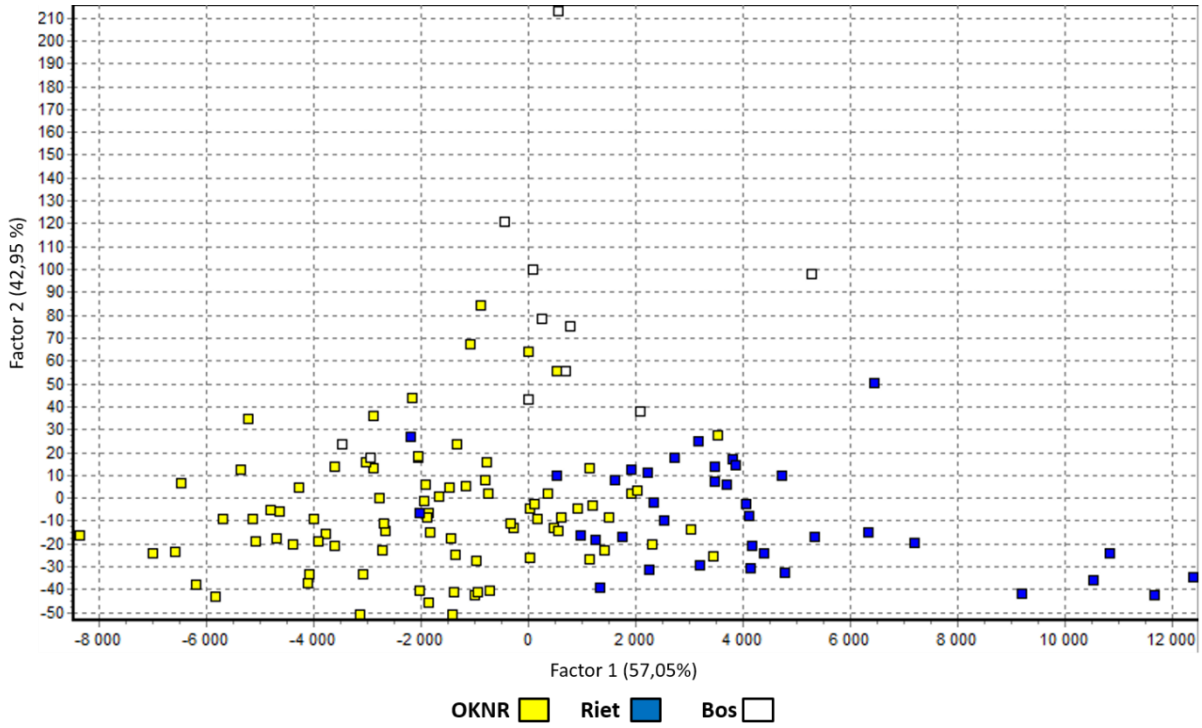


Figure 3.3: Factorial correspondence analysis showing scatter plots of the individual genotypes obtained using six microsatellite loci. Three clusters are visible based on the tight grouping of the individuals of the same population as represented by their colour. The yellow cluster is made up of the individuals belonging to the OKNR, whilst the blue is from Riet and white represents Bos.

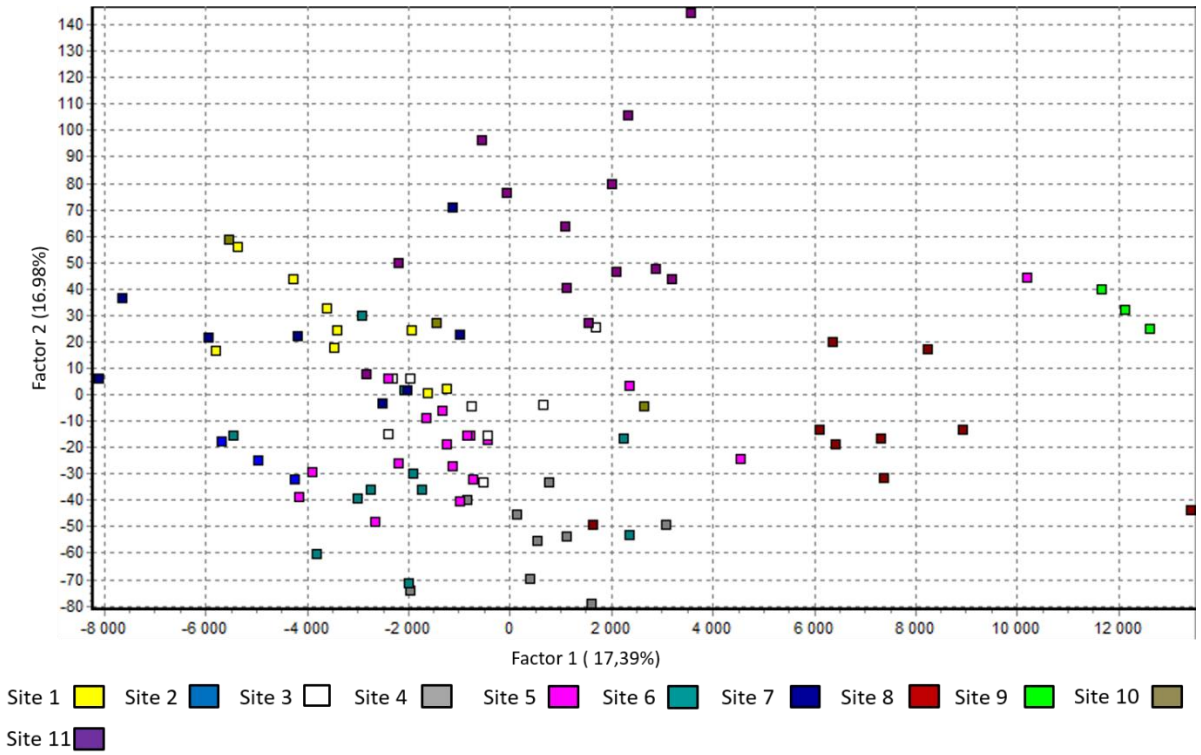


Figure 3.4: Factorial correspondence analysis showing scatter plots of the individual genotypes obtained using 6 microsatellite loci. Each colour represents the individuals of each of the 11 sampling sites within the OKNR. No clear clustering is visible among the 11 sites.

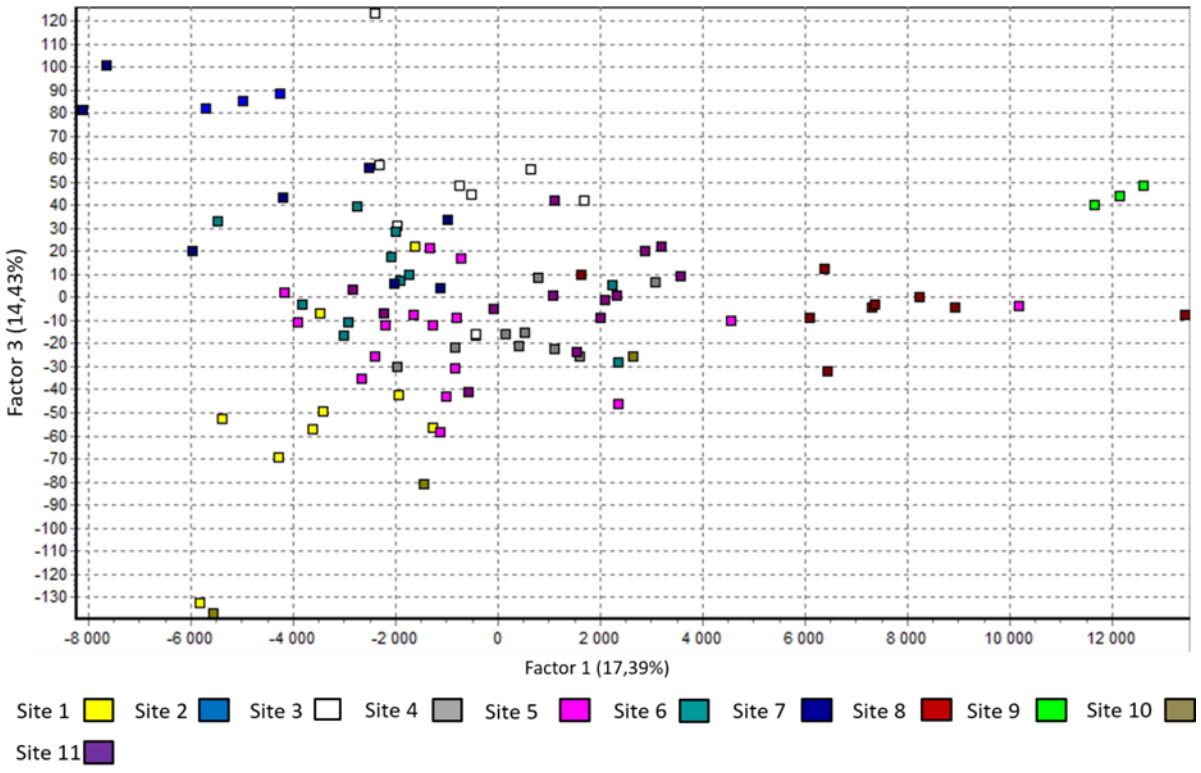


Figure 3.5: Factorial correspondence analysis showing scatter plots of the individual genotypes obtained using 6 microsatellite loci. Each colour represents the individuals of each of the 11 sampling sites within the OKNR. No clear clustering is visible among the 11 sites.

For both instances (OKNR *vs*. RIET *vs*. BOS and 11 sites of OKNR) the Bayesian clustering analysis implemented in STRUCTURE in conjunction with the method of Evanno *et al*., 2005 (Δ*K* statistic) implemented in Structure Harvester, identified the most likely number of populations as *K*=2 (Figure 3.6 A and B). (Figure 3.6 C and D) represent the proportional assignment of each individual to the proposed clusters, whilst grouped to its sampling population. Both structure plots depict no clear structuring between the various populations, thereby supporting the AMOVA in claiming very little to no differentiation of the populations. The Δ*K* statistics do, however also support the pairwise F_{ST}'s in identifying the number of clusters as 2 in each case, OKNR and Riet differentiated and Site 11 being different from some of the other sites.

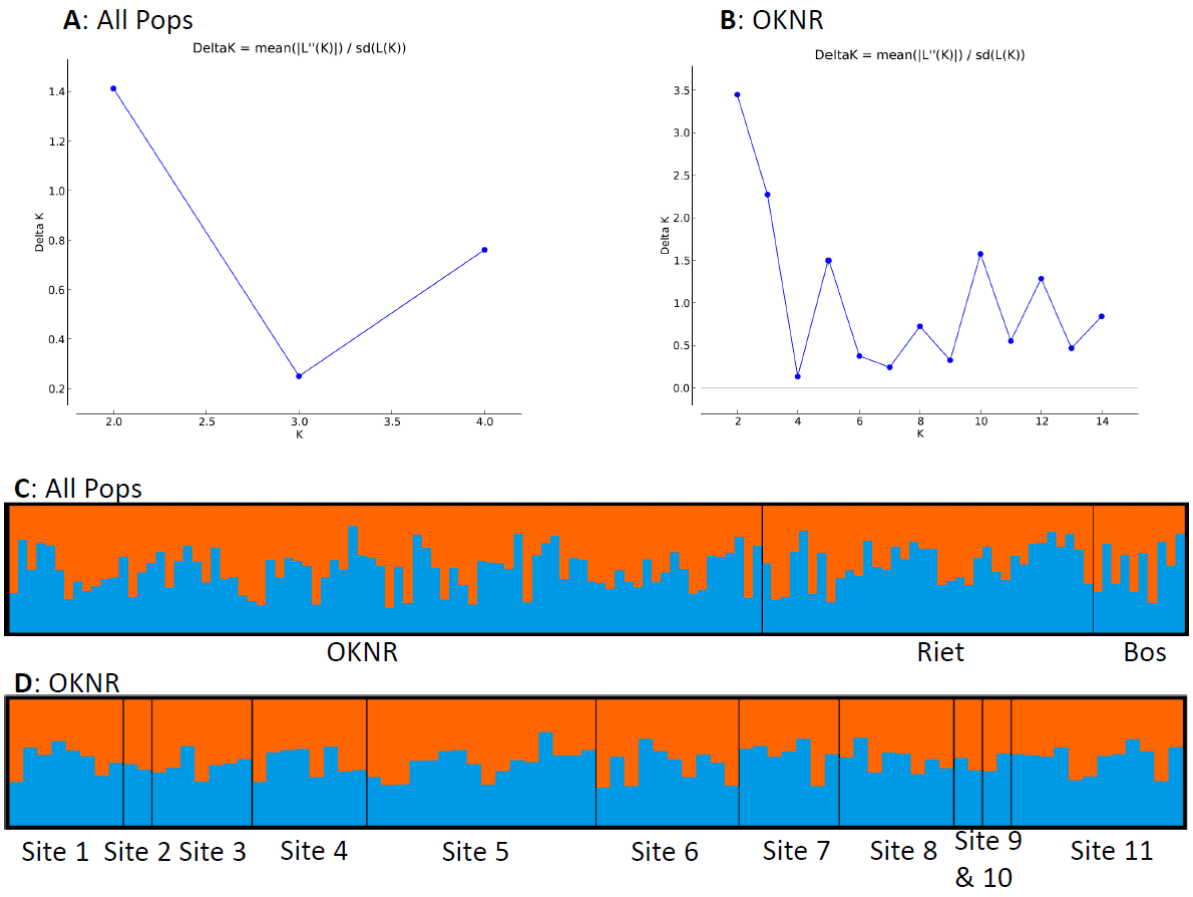


Figure 3.6 A) Δ*K* as a function of *K* following Evanno *et al*., 2005 for OKNR vs Riet vs Bos. B) Δ*K* as a function of K following Evanno *et al*., 2005 for the 11 sites within the OKNR. C) Proportion of model base clusters (*K*=2) in the ancestry of three populations

3.4.4 Microsatellite Data Analysis: Effective population size, N_e and population bottlenecks:

Effective population size (N_e) estimates using the Linkage Disequilibrium method indicated very low effective population sizes for the lower 95% confidence interval (95%CI) for Bos and Riet, whilst that of the OKNR was distinctly higher (Table 3.7). The point estimates using the Linkage Disequilibrium method for all three sampling regions were infinite. The heterozygosity excess method yielded no results of importance.

Table 3.7 N_e estimates for OKNR, Riet and Bos as calculated by the linkage disequilibrium and heterozygosity excess methods

Population	Methods and Estimates [95% CI]	
	Linkage Disequilibrium	Heterozygosity excess
OKNR	∞ [464.9 - ∞]	∞[∞ -∞]
Riet	∞ [118.4 - ∞]	∞[∞ - ∞]
Bos	∞ [14.6 - ∞]	∞ [38.5 - ∞]

The infinite allele model (IAM) shows significant heterozygosity excess ($P \leq 0.05$) for the OKNR and Riet populations, indicating a possible recent bottleneck in each of these populations. This is however not supported by the SMM and TPM models ($P > 0.05$). The Bos population showed no indication of a recent bottleneck event for any of the models tested. Overall, the alleles showed a normal, L-shaped distribution across all populations indicating the lack of sufficient evidence for population bottlenecks (Table 3.8)

Table 3.8 Bottleneck analysis of the 3 populations OKNR, Riet and Bos

Population	Model						Mode Shift
	IAM		TPM		SMM		
	H exc	H def	H exc	H def	H exc	H def	
OKNR	0.02344	0.98438	0.57813	0.50000	0.97656	0.03906	normal
Riet	0.01563	0.99219	0.07813	0.94531	0.94531	0.07813	normal
Bos	0.50000	0.57813	0.92188	0.21875	0.94531	0.07813	normal

The Wilcoxin signed rank test was used to test for heterozygosity excess or deficiency under the infinite alleles model (IAM), two-phase model (TPM) and the stepwise mutation model (SMM) *P*-values are indicated under each mutation model type with significance at $P \leq 0.05$. Parameters for TPM were: variance = 30; proportion of SMM = 70%. Estimates were based on 10 000 replications.

3.4.5 Microsatellite Data Analysis: Relatedness, *r*, per population:

Mean relatedness *r*, were non-significant for all three populations as they fall within the 95% CI.

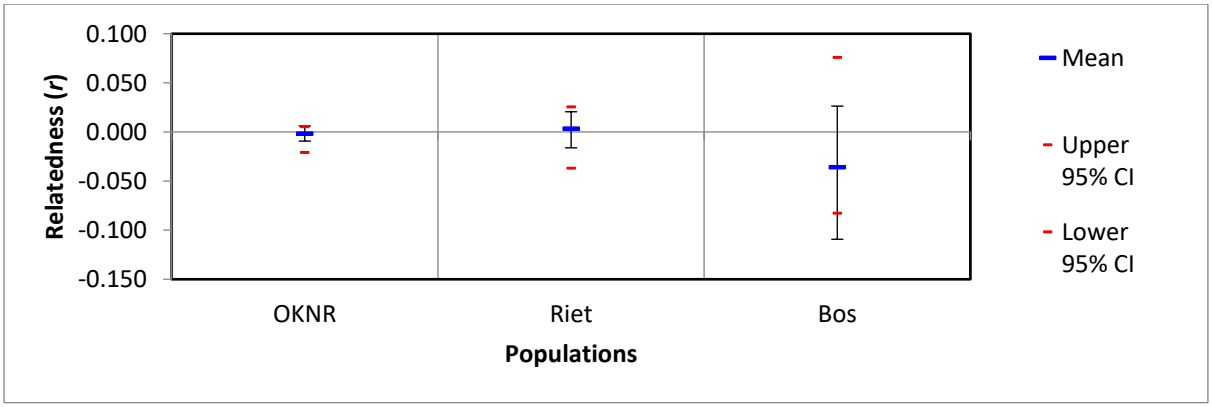
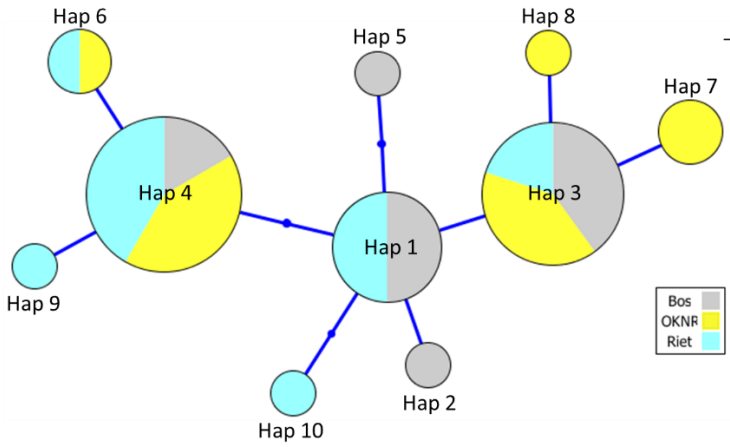


Figure 3.7 Estimates of mean relatedness per study population

3.4.6 Molecular diversity estimates and population structuring (Mitochondrial DNA):

A 736bp fragment of the mitochondrial control region (*D-Loop*) was successfully amplified and sequenced for 37 of the 39 (13 OKNR, 13 Riet and 11 Bos) individuals, whilst 2 individuals

from the Bos population failed to yield successful amplification. Analysis revealed 10 distinct haplotypes (Figure 3.8), comprising of three major haplotypes (Hap 1; Hap 3 and Hap 4) and an array of lower frequency variants stemming from each. The three major haplotypes are prevalent among the various sampling population although notably haplotype 1 (Hap 1) was absent from the OKNR population. The lower frequency variants seemed to be unique haplotypes, each to a specific population (Hap 2; Hap 5; Hap 7-10).



Haplotypes	Individuals
Hap 1(6)	Bos_4; Bos_8; Bos_11; Riet_21; Riet_34; Riet_38
Hap 2(1)	Bos_5
Hap 3(10)	Bos_6 – 7; Bos_9; Bos_16; OKNR_2; OKNR_38; OKNR_43; OKNR_50; Riet_29; Riet_33
Hap 4(12)	Bos_13 – 14; OKNR_3; OKNR_15; OKNR_24; OKNR_34; OKNR_78; Riet_1 – 2; Riet_30 – 31; Riet_36
Hap 5(1)	Bos_15
Hap 6(2)	OKNR_20; Riet_35
Hap 7(2)	OKNR_29; OKNR_57
Hap 8(1)	OKNR_64
Hap 9(1)	Riet_10
Hap10(1)	Riet_13

Figure 3.8 Median-joining network of *L. seeberi* mtDNA *D-loop* haplotypes. Haplotypes are separated by the blue branch lengths, with the basic branch indicating a single mutation and a branch with a notch in the middle indicating 2 mutational steps. The size of the circles are proportional to the frequency of the haplotypes. See the legend for sample numbers.

Of the 10 haplotypes illustrated in (Figure 3.8), Bos shares 5 haplotypes with two of these haplotypes unique to this region, Riet shares 6 of the haplotypes (2 unique haplotypes) and OKNR shares 5 haplotypes (2 unique haplotypes). Over all the populations, the haplotype diversity (h) and nucleotide diversity (π) ranged from 0,782 (OKNR) to 0,821 (Riet) and 0,00232 (Bos) to 0,00300 (OKNR) respectively (Table 3.9), indicating that genetic diversity across the populations is relatively high and in most part comparable.

Table 3.9 Summary of population diversity statistics for *Labeo seeberi* over all mtDNA *D-loop* haplotypes from each sampling location. *n*, number of samples; N_h, number of haplotypes (unique haplotypes); *h*, haplotype diversity; π, nucleotide diversity.

Population	*n*	N_h	*h*	π
OKNR	13	5(2)	0.782	0.003
Riet	13	6(2)	0.821	0.003
Bos	11	5(2)	0.818	0.002

Table 3.10 Analysis of Molecular Variance of *Labeo seeberi* across its three populations namely Bos vs. OKNR vs. Riet.

Source of variation	*Sum of squares*	*Variance components*	*% of Variation*
Among populations	*3.155*	*0.047*	*4.55*
Within populations	*285.124*	*0.994*	*95.45*
Total	*33.809*	*2.105*	
Φ_{ST}*: 0.045*	***P**: 0.152*		

***Statistical significance at the 5% nominal level (P ≤ 0.05)**

Analysis of Molecular Variance (AMOVA) results indicated no significant genetic divergences between the three populations (Φ_{ST} = 0.045, *P* = 0.152) making up for merely 4.55% of the total variance (Table 3.10). The largest component of variance was the diversity of haplotypes within each population contributing to 94.45% of the total variance (Table 3.10).

Table 3.11 Pairwise *ΦST* values describing the extent of genetic differentiation between 3 regions within the Olifants-Doring River system (Bos Riet and OKNR)

	OKNR	Riet	Bos
Riet	0.022	0.000	
Bos	0.026	0.094	0.000

*** Statistical significance at the 5% nominal level (P ≤ 0.05)**

Pairwise Φ_{ST} analyses revealed no sinificant genetic differentiation between the populations (Table 3.11). Of note, however, is the comparison between Bos and Riet as the differentiation between them is close to significant (Φ_{ST} = 0.094; *P* = 0.065) (Addendum H).

3.5 Discussion:

Genetic diversity as represented by number of alleles (A_N), allelic richness (A_R) and heterozygosity levels (Addendumm E) (Figure 3.2) are moderate to low as has previously been reported by Sahoo *et al*., 2014 and Singh *et al*., 2012 in their studies done on *Labeo rohita* and *Labeo calbasu* respectively. OKNR has the highest mean number of alleles, but very few private alleles (A_{PR}). This could indicate the gene flow from OKNR to Riet and Bos as it shared most of what would be its private alleles. The unique alleles at that are maintained in the OKNR could be low frequency alleles that have a low probability of being passed on downstream. This could make sense as OKNR is the largest and most well kept (pristine) of the three populations, thereby potentially maintaining much of its historical genetic diversity and thus having the most number of alleles (A_N) by maintaining low frequency alleles. It also serves as the main breeding grounds for the rest of the river system thereby sharing a lot of its genetic makeup (high frequency alleles) with populations downstream (Riet and Bos). This is further substantiated by the mean A_R for all three sampling regions being so comparable. This is as result of a subset of high frequency alleles being shared and maintained among all three sampling regions. This is also represented in the large amount of overlap between the Bos Riet and OKNR in the FCA-Plot (Fig Riet and Bos (the two smaller sampling regions believed to be fed solely from OKNR (Paxton *et al*., 2012), however have their own A_{PR} indicating that gene flow does occur from OKNR to Riet and Bos respectively, but not from Bos or Riet to OKNR. This again makes logical sense as OKNR is separated from the rest of the river system by a waterfall as is mentioned in Lubbe *et al*., 2015. The mitochondrial data also supports this movement of fish from OKNR to Riet and Bos, but not from either Bos or Riet to OKNR. This is most notable with Hap1 being well represented in Riet and Bos but completely absent in OKNR (Figure 3.8). The presence of private alleles (Figure 3.2) and private haplotypes (Figure 3.8) of Bos and Riet indicates that the OKNR may not be the sole contributor of *Labeo seeberi* in the river system and leads to reason that there is at least a second breeding ground for these fish. This is supported by a spawning event in the Biedouw River, but the survival of these fish could not be determined and therefore deemed an insignificant contribution (Lubbe *et al.*, 2015). The ΔK statistic for determining the most likely number of clusters between OKNR, Bos and Riet is $K = 2$ indicating two ancestral populations

being propagated (Figure 3.6). This is echoed in the results of the FCA-plot (Figure 3.3) and the F_{ST} values in (Table 3.3). The unequal genetic contribution of the OKNR population to genetic makeup of the system though largely overshadows that of the second ancestral population and therefore the large amount of genetic overlap between the sampling regions OKNR, Bos and Riet (Figure 3.2; Figure 3.3; Figure 3.6C and Table 3.5).

The mean heterozygosity for each of the sampling regions were relatively high in conjunction with low values for the Fixation index (F_{IS}) indicating no heterozygote loss and maintaining the genetic diversity available. There was also no indication of significant levels of relatedness (r) in any of the three populations. It would seem as although OKNR, Riet and Bos have moderate to low diversity, they are effective at maintaining that diversity and therefore healthy populations that are not inbred at this point in time. Although the effective population size is of concern, at least there is no indication of a recent bottleneck event indicating that although genetic diversity is moderate to low, at least it is being maintained. The low effective population size remains of concern though as critical minimum (>1000) is considered to be necessary to maintain genetic diversity and avoid the increase in frequency of deleterious alleles by inbreeding (Frankham *et al*., 2003; Palstra and Ruzzante, 2008; Waples and Do, 2010). Of more concern is Bos having N_e-estimates smaller than 50, which is deemed to be the point at which populations are at risk of inbreeding depression (Franklin, 1980). It is worth noting that the point estimates of each of the estimates are indicated as infinite, and is a result of the microsatellite markers and small population size not granting enough statistical power to accurately make predictions (Waples, 1989).

Results for the 11 sites of the OKNR indicate that there is no real significant sub structuring of the Sampling region OKNR as a whole and therefore can be seen as a single sampling region and managed accordingly.

3.6 Conclusion:

The presented data indicates that OKNR (the main breeding ground for *Labeo seeberi*) has no sub-structuring along its length and can be managed as a single unit. It also determines that there is gene flow from OKNR to Bos and Riet by means of migration, but no gene flow back

to OKNR. This is as result of an instream barrier in the form of a waterfall at the lower border of OKNR. There is also evidence for breeding grounds other than OKNR, which means that conservation efforts will need to be broadened from just OKNR to other sites to ensure genetic diversity is maintained. This study recommends that the fish of the OKNR be managed as a separate management unit (MU) to that of Riet and Bos. Since there was no structure between the 11 sampling sites of the OKNR, This region can be managed as a single Management Unit (MU) and does not need to be further subdivided. As Bos and Riet populations are impacted by migration from OKNR but do not themselves contribute genetically to the OKNR and also contain their own unique diversity as represented by the unique alleles we propose the populations of the Mainstream Doring River be managed as separate MU. As there is very little differentiation between populations and OKNR contributing to most of the diversity, OKNR, Bos and Riet can be managed as a single evolutionary significant unit (ESU), but with strict guidelines indicating no translocation of fish from Riet or Bos to OKNR

Chapter 4: Concluding remarks and future studies:

4.1 Overview of research findings:

The protection of *Labeo seeberi* in its last remaining natural habitat is important not only for the ecological role it fulfils, but also for the scientific knowledge to be gained and applied to other conservation efforts. This is especially true as no current conservation efforts extend to ensure the successful recruitment and sustainability of *Labeo seeberi* (Jordaan, Impson and van der Walt, 2011; Paxton *et al*., 2012; Lubbe *et al*., 2015).

This has encouraged the drafting of a Biological Management Plan for the Clanwilliam sandfish by Paxton *et al*., *2012*. One of the key points of this BMP-S is to increase the biological information on this species in order to create a successful and sustainable conservation programme. This is especially true for genetic information regarding the species as, very few genetic studies and no population genetic studies have been done on this species to date (Paxton *et al*., 2012; Jordaan, Impson and van der Walt, 2011). In this study the objectives were to increase the biological information by inferring life history traits from a close living relative species by means of common ancestry (species relatedness) using the *CO1* mitochondrial gene. Also, to quantify genetic diversity and population genetic structure of *Labeo seeberi* in its last remaining natural distribution using the multi-locus data generated from six microsatellite loci and the mitochondrial control region (*D-loop*).

Effective management strategies have evolved significantly in the last two decades, becoming multi-faceted comprehensive studies to delineate as many factors playing a role in the ecology of a species as possible (Scribner *et al*., 2016). This includes monitoring invasive species, water level fluctuations (seasonal changes, water abstraction or obstruction etc.), behaviour, life history traits and population genetics. Understanding the population dynamics and structure of a species give the researcher insight into the general genetic welfare of the species and their risk of extinction as measure of their genetic diversity (or lack thereof) and by extension effective population size. It also informs the researcher on distinct stocks, or evolutionary significant units, which in turn helps the effective management of these stocks and aids restocking programmes. Using mitochondrial and microsatellite markers grant

researchers the opportunity to gain valuable insight into the population genetics and genetic structure of the species in question (Vrijenhoek, 1998; Scribner *et al*., 2016).

In an effort to gain more knowledge of the life history traits of *Labeo seeberi*, Chapter 2 presented the closest evolutionary relative to the species. In short, all available *Labeo* species on the public repositories (NCBI, EMBL etc.) were tested for evolutionary relatedness using the *CO1* and *Cytb* mitochondrial genes. This was done by constructing species relatedness trees using Neighbor-Joining and Maximum Likelihood algorithms (and best fitting substitution models). The result was an alignment of 478bp fragments comprising of 34 species for *CO1* and 275 bp fragments comprising of 24 species for *Cytb*. The *CO1* trees for both NJ an ML displayed a distinct separation between African species and Asian species. For *Cytb* the NJ and ML trees showed no distinction between African and Asian *Labeo* species. Cytochrome b was thus deemed to not be a successful mtDNA region in inferring ancestry between *Labeo* species. Using the results of *CO1* based off this study alone (Figure 2.2 and Figure 2.3), it is speculated that *Labeo vulgaris* (Nile carp) is most closely related to *Labeo seeberi* and thereafter *Labeo horie*. Unfortunately, not much is known about *Labeo horie* so no comparisons to *Labeo seeberi* could be made. It is important to note that the other eight *Labeo* species from South Africa were omitted in this analyses as the sequences were unavailable. This study did however speculatively support the proposed groupings of Reid, 1985 for *Labeo* species in Africa, namely *Labeo niloticus* group (LNG), *Labeo forskalii* group (LFG) and *Labeo coubie* group (LCG) as represented in Figures 2.2 and 2.3. Following on from this trend, it would be possible to assume that *Labeo seeberi* would group with the *Labeo umbratus* group (LUG) as proposed by as determide in the study of Ramoejane, 2016 which included all nine South African *Labeo* species. The LUG is comprised of *Labeo seeberi*, *Labeo umbratus*, *Labeo capensis* and *Labeo rubromaculatus*. Based on morphology and distribution, *Labeo capensis* was deemed the closest relative to sandfish (Ramoejane, 2016). *Labeo capensis* is also similar in size to *Labeo seeberi*, spawns during the same time and has a very similar habitat. Using *Labeo capensis as reference* we could approximate the growth rate of *Labeo seeberi* to be 4-6cm per year for the first 6 years where after growth gradually slowed down. We also could infer the time to maturity as ±4years old (Baird, 1976). Knowing the rate of growth researchers can estimate the time which it takes for juveniles to grow beyond the prey size class (250mm) at which it is safe to return to the mainstream Doring River. Age to

maturity also gives good insight into planning breeding programs, understanding generation time and other analysis that are dependent on identifying distinct generations. This has greatly helped in gaining of preliminary life history traits in this critically endangered species without further having to sacrifice individuals for study.

In Chapter 3, six microsatellite loci and a mitochondrial marker were used to investigate population genetic structure and historical demographics of *Labeo seeberi* in the Olifants-Doring River system. For the microsatellite analysis: The per population genetic diversity summary statistics, in conjunction with the factorial correspondence analysis (FCA), Analysis of Molecular Variance (AMOVA) and Bayesian clustering analysis indicate very shallow genetic differentiation between the three sampling populations (OKNR, Riet and Bos). As for the 11 sampling sites within the OKNR, there was no indication of structure and the OKNR "sanctuary" population deemed panmictic. The genetic diversity summary statistics also hint at a bottleneck event. Testing for a bottleneck, however, indicated that it must have happened more than 4x generations prior to the current as no contemporary bottleneck was evident. This is supported by survey data indicating that the largest decline in population number and distribution happened during the 1950's, much more than four generations ago (Paxton *et al*., 2012). Effective population sizes for all three sampled populations are of concern as they could be rather low (based on the lower confidence bound estimate) and thus under threat of extinction. Reduction in population size is often associated with a decrease in genetic diversity. Reduced genetic diversity may result in decreased population viability and an increased extinction likelihood (number of alleles and allelic diversity of a population, dictate the populations adaptability to changing environments.). This is especially true for threatened or endangered species than closely related non-threatened taxa (Spielman *et al*., 2004; Markert *et al*., 2010; Willoughby *et al*., 2015). Currently, the sampling populations seem healthy, with no significant relatedness (*r*) detected and the F_{IS}-values for all three populations being very low (F_{IS} < 0.1) indicating very little inbreeding in these populations. For the mitochondrial analysis: the weak population differentiation between OKNR, Riet and Bos was echoed by F_{ST} results. Diversity estimates and haplotype network did however indicate a moderate level of haplotype diversity. This is good as it means more than one maternal lineage is propagating, thereby maintaining diversity in the population. What is of note though is the number of unique haplotypes present outside of the OKNR. This means

that the OKNR is not the sole contributor of offspring as is implied by recent literature (Lubbe *et al*., 2015), making a strong case for the conservation of populations outside of the OKNR in order to maintain as much diversity as possible. *Labeo seeberi*, although few in number and sporadic in distribution, at least for the time being are healthy with whatever small rate of migration maintaining diversity and keeping population differentiation to a minimum.

4.2 Biological significance of research findings:

4.2.1 Evolutionary relatedness:

A common problem for working on scarce endangered species with limited distribution is the lack of scientific information about them. This is a significant hurdle in planning multifaceted conservation programmes as many of the parameters are unknown, thereby effecting the efficacy of such a programme (Paxton *et al*., 2012; Lubbe *et al*., 2015; Scribner *et al*., 2016). One such facet of importance is the life history traits of the species, such as preferred habitat, feeding, growth rate, time to maturity and behaviour. The lack of research done on Clanwilliam sandfish is very apparent and therefore shares many of the shortcomings listed above. This study provides some speculative preliminary insight into the closest evolutionary relatives (that were available to this study) to sandfish using mitochondrial genes. This allowed us to preliminarily identify the closest relative of *Labeo seeberi* to be *Labeo vulgaris* (Figure 2.2 and 2.3), based off our own data (Excluding the other eight *Labeo* species in South Africa as in Ramoejane, 2016). *Labeo vulgaris* inhabits the freshwater rivers of north-eastern Africa (Egypt, Ethiopia and Sudan). *Labeo vulgaris* is benthic feeding species just like *Labeo seeberi* and spawns in the fast flowing waters of tributaries During May to June (Spring in the Northern Hemisphere). Sexual maturity for *L. vulgaris* is reached around two years old (220mm-260mm TL) and grow to ±470mm maximum length (Azeroual *et al*., 2010). *Labeo vulgaris* closely resembles *Labeo horie* based on morphology. Furthermore, *Labeo senegalensis* and *labeo horie* pair very close together, confirming Yang *et al*., 2012 in that they are sister species. The grouping of *Labeo horie, Labeo senegalensis, Labeo altivelis* and *Labeo weeksii* mirror that of the proposed *Labeo niloticus* group (LNG) by Reid, 1985. The grouping

of fish species on Figure 2.2 and 2.3 also represent the *Labeo forskalii* group LFG and *Labeo coubie* group (LCG). Extrapolating from the findings of Ramoejane 2016 (based of the work of Lowenstein *et al*., 2011 and Yang *et al*., 2012) *Labeo seeberi* belongs to the *Labeo umbratus* group (LUG). We can now speculatively infer *Labeo capensis* as the closest living relative to *Labeo seeberi*, not only geographically, but also genetically and they share similar habitats (Ramoejane, 2016). Using this information and other anecdotal evidence and observations, *Labeo seeberi* is rheophilic, adapted to fast flowing waters or tributaries of large river systems. Migrate from feeding to spawning grounds during September – November (Spring in the southern Hemisphere) and are benthic feeders. Further inferring biological traits from *Labeo capensis* we speculate *L. seeberi* to reach sexual maturity around 4 years old (250mm TL) and has an approximate growth rate is 40-60mm per year up to year six where after it decreases in rate steadily. *Labeo seeberi* reaches a maximum size of 650mm (Baird, 1976 ; Paxton *et al*., 2012; Lubbe *et al*., 2015). Using this preliminary speculative information, authorities can plan more accurate initial conservation programmes until further studies provide definitive results.

4.2.2 Population Genetics of Clanwilliam sandfish:

The cumulative threats posed by the Olifants-Doring River system in its current state have seen a decline in both number and distribution in almost all of the endemic fish species of the region (Jordaan, Impson and van der Walt, 2011). These threats include the variable water levels throughout the year reducing vast water-bodies to become mere streams or completely dried-up (as result of the arid nature of the area during summer months, water abstraction by farmers and water infrastructure such as dams), predatory invasive fish species and the destruction of riparian zones. This causes the endemic populations to be severely fragmented, thereby reducing the population sizes and potentially serving as barrier to gene flow (Jordaan, Impson and van der Walt, 2011; Paxton *et al*., 2012; Lubbe *et al*., 2015). Small isolated populations are in turn more vulnerable to genetic drift or inbreeding (Wright, 1931; Palstra, 2008; Hare *et al*., 2011). If unchecked, these could reduce genetic diversity and therefore adaptive potential as well as decrease fitness (Willi, 2006; Hare *et al*., 2011). The outcome is that these small populations if left without the balancing effects of gene flow become all the

more susceptible to extinction over time (Lynch and Lande, 1998; Charlesworth and Willis, 2009; Frankham, 2005). This study provides some insight into the population genetics of these disjunct populations of the Clanwilliam sandfish. The assessment revealed that the effective population sizes (N_e) of the various sampling populations varied greatly (Table 3.7). The Bos population is of most concern as it indicated a $N_e < 50$ for the lower 95% confidence interval for both the linkage disequilibrium and heterozygosity excess methods (14.6 and 38.5) respectively. Effective population size of <50 indicate that the population is partially isolated and at risk of losing genetic diversity at a much increased rate as compared to a panmictic population (Lacey and Lindenmayer, 1995). For the Riet population $50 < N_e > 500$ therefore less critical than the Bos population, but still at risk of losing genetic diversity (adaptive potential) in the long run. It is worth noting that Riet is on the lower end of 500 ($N_e = 118$) and therefore still at some risk of rapid loss of genetic diversity. Effective population size < 500 are at risk of losing genetic diversity in the long term as compared to panmictic population (Lande, 1995; Lacey, 1997). It appeared that some moderate level of gene flow maintained the genetic diversity and counteracted the effects of inbreeding as the relatedness (r) of individuals within each population was deemed to be non-significant (Figure 3.7) as well as very low F_{IS}- values (< 0.1) (Figure 3.2). This was further supported by the genetic structure between populations being very shallow as represented in the pairwise F_{ST}-values (Table 3.3) and FCA-plot (Figure 3.3). This is echoed by a study done by Palstra and Ruzzante 2008 in which they found a decrease in effective population size to cause an increase in risk of extinction, especially if these small populations are genetically isolated. The counterpoint, however was that if gene flow by migration was present, its effects will increase as the population size got smaller. Thus, only a few individuals need to migrate to these semi isolated populations to maintain genetic diversity and thereby fitness (Ellstrand and Elam, 1993; Vrijenhoek, 1998). The populations although fractured thus seem to be "genetically" healthy and non-distinct. In the context of the Olifants-Doring River System this makes sense as the OKNR is the largest and healthiest population (species sanctuary) that feeds both Bos and Riet populations, thereby maintaining genetic diversity. The threat of extinction though still remains as many of the alleles maintaining diversity are in low frequency and thus subject to loss by genetic drift if gene flow were to stop (Vrijenhoek, 1998; Hare *et al*., 2011). The populations remain small and vulnerable to disease, inbreeding or extinction by natural disaster (Wright, 1931; Charlesworth and Willis, 2009; Palstra, 2009; Hare *et al*., 2011). One

last thing to note was the private alleles (Figure 3.2) and haplotypes (Figure 3.8) of populations outside of the OKNR indicating that the OKNR is not the sole contributor of offspring (successful recruitment) in the Olifants-Doring River system and thus warrants the extension of conservation to populations outside the OKNR in order to maintain diversity.

4.2.3 Conservation implications:

Labeo seeberi plays an integral role in controlling algae levels and cycling nutrients in the native rivers in which they occur, keeping the river and ecosystem healthy for the other species who share it with *Labeo seeberi* (Paxton et *al*., 2012; Ramoejane, 2016). The drastic decline of sandfish in its natural habitat over the last 50 years and the continuation of this trend has called for conservation action in order to protect and maintain this species (Paxton *et al.*, 2012; Lubbe *et al.*, 2015). In Chapter 3 this study found the populations to be genetically healthy, the population distribution is fractured, with effective population sizes of the Riet and Bos populations remaining low significantly less than 500 and therefore under threat for not being able to reliably sustain these populations for many generations to come (Lande, 1995; Lacey, 1997; Frankham *et al*., 2003; Palstra and Ruzzante, 2008). Riet and Bos populations are being kept healthy by migration and gene flow from OKNR to either of these populations, maintaining genetic diversity, including the unique alleles present in Riet and Bos. It is thus critical that the movement of fish from OKNR to the rest of the river system is not cut-off (Vrijenhoek, 1998; Palstra and Ruzzante, 2008). Migration from Riet and Bos to OKNR is, however not possible as highlighted in Chapter 1 by the waterfall that makes up the southern border of the OKNR and serves as barrier against the rest of the River System making it ideal as a sanctuary (Paxton *et al*., 2012; Lubbe *et al*., 2015). This study also found a great deal of native diversity outside of the OKNR, thereby warranting the protection of populations outside of the OKNR to maintain the genetic diversity. As written by Coates *et al*., 2018, the goal of a successful conservation programme should be to maintain the dynamic processes, such as natural selection, gene flow and genetic drift that shaped the diversity within and between populations. This innate diversity may affect a species ability to react to changing environments, disease etc. It is therefore vital for conservation efforts to extend beyond the OKNR and ensure greater successful recruitment of juveniles in order to increase the effective

population size, reduce the impact of predation and re-establish recruitment sites and gene flow. Successful recruitment of juveniles, lessening the impact of predation and the possible re-establishment of recruitment sites by understanding the life history and biology of *Labeo seeberi* was explored in Chapter 2. This was done by inferring these traits from speculated closely related relatives. Knowing the time to maturity gives conservationist and or breeders a clear indication of generation time and the lead time required for a restocking project as it will take at minimum four years (Baird, 1976; Paxton *et al*., 2012; Lubbe *et al*., 2015). Growth rate is also important to gauge when a fish will be beyond the prey size class (>250mm) and is ready to be re-introduced to the wild or translocated (Baird, 1976).

Based on the results of Chapter 2 and 3 it is evident that conservation initiatives need to be put in place in order to secure future generations of *Labeo seeberi*. One of the principle components of a conservation effort is the establishment of a breeding programme in order to secure juvenile recruitment (Paxton *et al*., 2012). Two options are available, each with pros and cons. The first is captive breeding programmes. This entails breeding fish in captivity, in an artificial environment where most of the variables can be controlled. Pros to this method include a constant food supply, no invariability of environmental conditions, fish can be treated for diseases and the fish are easier to identify and work on. There is also no threat of predation. Captive breeding thus has a much higher recruitment of juveniles during spawning than other methods (if the fish respond to spawning signals). Cons are the limited broodstock used to start the breeding programme as studies show that small founding populations quickly lose most of their rare alleles (Allendorf, 1986; Leberg, 1991). This could be as result of unequal broodstock contributions, skewing the allele frequencies in the following generations. Careful genetic management is thus required in establishing the broodstock population to avoid genetic bottlenecks (Lacey, 1989; Vrijenhoek, 1998). Another way to combat the loss of genetic diversity in a captive breeding programme is to periodically introduce genes from wild stock (Vrijenhoek, 1998). Further cons are that captive bred fish are not acclimated to the environmental changes when introduced to the wild and can succumb to the natural stresses such as disease and predators, thereby negating the purpose of the programme. To combat this, captive bred fish can be put through an acclimatisation period to get use to the wild in a safe environment before being introduced fully. The second is natural breeding system where reserves or protected areas are created in the fish natural

habitat, to promote successful recruitment of juveniles. Pros are that the fish are acclimated to the environment, predators, and possible diseases. The fish also respond to natural spawning signals and don't run the risk of not spawning in captivity. Another advantage is that these protected areas can be population specific, thereby maintaining diverged populations and their adaptations as opposed to having an off-site breeding facility that distributes fish to all the sites thereby homogenising the different populations (Vrijenhoek, 1998). Disadvantages are that environmental variables are not controllable, fish are subject to disease and predation and the overall recruitment of juveniles is lower. Currently the BMP-S for Clanwilliam sandfish makes use of a natural breeding system where the OKNR serves as the protected area for recruitment of juveniles (Paxton *et al*., 2012; Lubbe *et al*., 2015). However, this does not secure the rare alleles and haplotypes of the Riet and Bos populations. Creating protected zones at either Riet or Bos will retain these haplotypes and offer a chance to extend the distribution range of *Labeo seeberi* throughout the Olifants-Doring River system. Protected zones also offer the chance to eradicate alien vegetation and fish species, reclaiming habitat for the native species (Weyl, 2014).

In conclusion, the Oorlogskloof Nature Reserve should be managed as a separate Management Unit (MU) to that of the populations downstream of the waterfall *i.e.* Bos and Riet. This is based on the fact that OKNR is isolated from gene flow to OKNR from the rest of the river system as result of the waterfall at its southern border, and is therefore functionally independent (Moritz, 1994; Palsboll *et al*., 2007). Gene flow from OKNR to Rietkuil and Bos must however be maintained but no translocation of fish should take place from Riet and Bos to OKNR. Bos and Riet then form the second MU with OKNR, Bos and Riet combined making a single Evolutionary Significant Unit (ESU) as OKNR contributes the majority of the genetic make-up of the species and, therefore the majority of the evolutionary heritage (Palsboll *et al*., 2006). Should the Bos-Riet MU manage to retain its rare alleles and haplotypes and increase in frequency in the future, it could be re-assessed to determine if it is an ESU.

4.3 Limitations and Future Perspectives.

The following limitations were present in this study. Sampling bias was a major limitation as all sampling locations did not have equal number of representatives. This is as result of the difference in abundance of fish in the different regions and the haphazard nature of the fishing exercise. This led the sample size to be heavily weighted toward the sandfish sanctuary, OKNR as it is home to the largest population of sandfish (Paxton *et al*., 2012; Lubbe *et al*., 2015). The sampling bias thus follows the curve of relative abundance. Future studies could definitely benefit by increasing the number of individuals sampled from these regions. Another limitation was the number of sampling sites/populations included in this study. The presented sampling sites do not represent the entire distribution of sandfish. Future studies could benefit from sampling more tributaries and getting a more comprehensive insight into the population genetics of sandfish across its entire distribution. This could be done by sampling during spawning season (September to November) while adult fish are migrating and then aggregating in tributaries in larger numbers. This would also more accurately display the individuals contributing to the gene pool of the next generation for each specific population. Increasing sample size per population would also yield greater statistical power to more accurately determine population genetic parameters and put researchers instead to look at certain historical demographic events in more detail (on a finer scale)(Puechmaille, 2016). The study to find the most recent common ancestor/most closely related species was also flawed in that none of the other eight *Labeo* species in South Africa were included as they are geographically the closest and expectedly genetically the closest. Future prospective would be to include these eight other species so to have a more complete representation of the *Labeo* species (Yang *et al*., 2012; Ramoejane, 2016). Another limitation was the mitochondrial gene used for the evolutionary relatedness study. Ideally, a minimum of four markers should be used for conducting such studies, as by increasing the number of genes, the more accurate the evolutionary history of the species becomes (Yang *et al*., 2012; Ramoejane, 2016). This then ties in with the future prospective of either using four genes to repeat this analysis or even better, sequence the entire mitochondrial genome (Note: This is subject to specimen availability or availability of complete mitochondrial sequences for *Labeo* species online. With the few number of samples obtained, another limitation was the few

number of microsatellite markers used in this study and the limited informativeness of these markers. Ideally a population genetic study would at minimum strive to have 10 informative microsatellite markers with the option of adding more markers if so permitted or able. This is because the statistical power of each of the analysis increase by the number of informative markers added (Landguth *et al*., 2012). Future studies could benefit from a genotyping by synthesis approach, which would result in a genome wide genotype of each individual incorporating the information of hundreds of microsatellite markers per individual across all individuals. The result is increase in statistical confidence and it also minimizes the microsatellite marker bias itself as these markers are across the entire genome. Alternatively, this technique could be applied to a single individual, the microsatellites identified, the best performing ones selected and that these microsatellite markers be used for downstream analysis (Robeldo, 2018). The upside of this is that the markers are species specific and that there are many markers to choose from. This makes it ideal as the researcher can pic markers that are very informative, amplify consistently and are easy to score. A future prospective would also be to have samples of discrete generations as to not bias estimates of relatedness and effective population size.

4.4 Conclusion:

Although *Labeo seeberi* numbers and distribution have declined, the populations that do persist seem to be stable with very little inbreeding and relatedness. This is maintained by migration and gene flow from the OKNR to the Riet and Bos populations, counteracting the effects of genetic drift. Riet and Bos do, however have unique haplotypes and alleles thereby warranting a Riet-Bos MU and an OKNR MU. This study also managed to infer much needed biological history of *Labeo seeberi* such as growth rate and time to maturity. It also identified close living relatives that could be used as reference for designing management plans or adapting the *Labeo seeberi* management plan for.

References:

Abdul-Muneer, P. M. 2014. Review Article: Application of Microsatellite Markers in conservation Genetics and Fisheries Management: Recent Advances in Population Structure Analysis and Conservation Strategies. Genetics Research International volume 14.

Abrahams, A.A.M. and W.A.J. Pretorius. 2000. The freshwater fish fauna of Oorlogskloof Nature Reserve, Northern Cape, South Africa. Unpublished Report. Northern Cape Nature Conservation Service, Kimberley.

Allendorf, F. W. 1986. Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biol.*, *5*, 181‑190. https://doi.org/10.1002/zoo.1430050212.

Azeroual, A., Getahun, A. & Vreven, E. 2010. *Labeo niloticus. The IUCN Red List of Threatened Species 2010*: e.T182692A7945029. http://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T182692A7945029.en

Bandelt H. J., P. Forster, A Rohl. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol.; 16(1):37–48. doi: 10.1093/oxfordjournals.molbev.a026036 PMID: 10331250.

Belkhir, K., P. Borsa, J. Goudet, L. Chikhi, F. Bonhomme. 2000. *GENETIX*, logical sous Window TM pour la génétique des populations. Laboratoire Génome et Populations. Montpellier.

Birkhead, A.L.; Boucher, C.; Brown, C.; Dollar, E.; Harding, W.; Kamish, W.; Paxton, B.R.; Permberton, C. and Ratcliffe, S. 2005. Olifants Doring Catchment Ecological Water Requirements Study. Riverine RDM Report. Volume 1: Specialist Reference Report. *Project No. 2002-376*. Southern Waters Ecological Research and Consulting for the Department of Water Affairs and Forestry: RDM Directorate. 309 pp.

Brownlie, S., C. De Villiers, A. Driver, N. Job, A. Von Hase and K. Maze. 2005. Systematic conservation planning in the Cape Floristic Region and Succulent Karoo, South Africa: enabling sound spatial planning and improved environmental assessment. Journal of Environmental Assessment Policy and Management Vol. 7(2): 201–228

Cambray, J.A.; King, J.M. and Bruwer, C. 1997. Spawning behaviour and early development of the Clanwilliam yellowfish (*Barbus capensis*; Cyprinidae), linked to experimental dam releases in the Olifants River, South Africa. *Regulated Rivers: Research & Management,* 13**:** 579-602.

Carrizo, S. F., Smith, K. G. & Darwall, W. R. T. (2013): Progress towards a global assessment of the status of freshwater fishes (Pisces) for the IUCN Red List: application to conservation programmes in zoos and aquariums. *International Zoo Yearbook* 47, 46-64.

Chakona, A., Swartz, E. R., Gouws, G. 2013 Evolutionary Drivers of Diversification and Distribution of a Southern Temperate Stream Fish Assemblage: Testing the Role of Historical Isolation and Spatial Range Expansion. PLoS ONE 8(8): e70953. doi:10.1371/journal.pone.0070953

Chakona, A., E. R. Swartz, G. Gouws and P. Bloomer. 2013. A freshwater fish defies ancient mountainranges and drainage divides: extrinsic and intrinsic influences on the evolutionary history of a recently identified galaxiid. J. Biogeogr 40: 1399-1412

Charlesworth, D., J. H. Willis. 2009. The genetics of inbreeding depression. Nature Reviews Genetics volume 10: 783 – 796

Coates, D. J., Byrne, M. and Moritz, C. 2018. Genetic Diversity and Conservation Units: Dealing With the Species-Population Continuum in the Age of Genomics. Front. Ecol. Evol. 6:165. doi: 10.3389/fevo.2018.00165

Consuegra, S., Verspoor, E., Knox, D. *et al*. 2005. Asymmetric gene flow and the evolutionary maintenance of genetic diversity in small, peripheral Atlantic salmon populations Conservation Genetics. 6: 823. https://doi.org/10.1007/s10592-005-9042-4

DAFF (Department of Agriculture, Forestry and Fisheries). 2012. *Status of the South African marine fishery resources.* Cape Town: Chief Directorate Fisheries Research, Fisheries Branch, Department of Agriculture, Forestry and Fisheries. p 60 – 64

Darwall, W. R. T., Smith, K. G., Tweddle, D. & Skelton, P. 2008. The status and distribution of freshwater biodiversity in southern Africa. *IUCN Red List of Threatened Species* 2008, 1-36.

Darwall, W. R. T., Smith, K. G., Tweddle, D. & Skelton, P. 2009. The status and distribution of freshwater biodiversity in southern Africa. *IUCN Red List of Threatened Species* 2009, 1-36.

De Moor, F. C., J. A. Day. 2013. Aquatic biodiversity in the mediterranean region of South Africa. Hydrobiologia 719(1): 237 - 268

De Moor, I. J., M. N., Bruton. 1988. Atlas of alien and translocated indigenous aquatic animals of southern Africa. S Afr Nat Sci Prog Rep No 144, CSIR, Pretoria, pp 78- 133

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J. and Ovenden, J.R. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. Mol Ecol Resour, 14: 209-214. doi:10.1111/1755-0998.12157

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z., Knowler, D. J., Leveque, C., Naiman, R. J., Prieur-Richard, A-H., Soto, D., Melanie L. J. Stiassny, M. L. J. & Sullivan, C. A. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews* 81, 163-182.

DWAF. 2005. Olifants/Doorn Water Management Area: Internal Strategic Perspective. *DWAF Report No. P WMA 17/000/00/0305*. Ninham Shand (Pty) Ltd in association with Jakoet and Associates, Umvoto Africa, FST and Tlou and Matji for the Department of Water Affairs and Forestry National Water Resource Planning, Pretoria.

Earl, D. A., and B. M. von Holdt, 2012 STRUCTURE HARVESTER: A website and program for visualising STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 2: 359-361 doi: 10.1007/s12686-011-9548-7.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 32(5):1792–1797. Published 2004 Mar 19. doi:10.1093/nar/gkh340

Ellstrand, N.C. and Elam, D.R. 1993 Population genetic consequences of small population size: Implications for plant conservation. Annual Review of Ecology, Evolution, and Systematics, 24, 217-242. http://dx.doi.org/10.1146/annurev.es.24.110193.001245

Evanno, G., Regnaut, S. And Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology, 14: 2611-2620. doi:10.1111/j.1365-294X.2005.02553.x

Excoffier, L., H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res. 10: 564 - 567.

Falush, D., M. Stephens, J. K. Pritchard. 2003. Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics 164: 1567 - 1587.

Frankham, R. 2015. Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. Molecular Ecology 24: 2610 – 2618.

Frankham, R., Bradshaw, C. J. A., & Brook, B. W. (2014). Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, *170*, 56‑63. https://doi.org/10.1016/j.biocon.2013.12.036.

Frankham, R. 2005. Genetics and extinction. *Biological Conservation*, **126**, 131–140.

Frankham, R., Ballou JD, Briscoe DA 2003. *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, UK. Franklin IR (1980) Evolutionary change in small populations. In: *Conservation Biology An Evolutionary-Ecological Perspective* (eds Soule ME, Wilcox BA), pp. 135–150. Sinauer, Sunderland, Massachusetts.

Fraser, D.J., Jones, M.W., McParland, T.L., Hutchings, J.A. 2007. Loss of historical immigration and the unsuccessful rehabilitation of extirpated salmon populations. *Conservation Genetics*, **8**, 527–546.

1914 Funk, W. C., Mckay, J. K., Hohenlohe, P. A., and Allendorf, F. W. 2012. Harnessing genomics
1915 for delineating conservation units. Trends Ecol. Evol. 27, 489–496. doi:
1916 10.1016/j.tree.2012.05.012

1917 Gaigher, C.M. 1973. Die status van inheemse vis in die Olifants Rivier. *Cape Nature*
1918 *Conservation Internal Report* Cape Town. 37 pp.

1919 Gaigher, I.G. 1984. Reproduction of *Labeo umbratus* (Pisces, Cyprinidae) in Wuras Dam, a
1920 shallow, turbid impoundment. *South African Journal of Zoology,* 19**:** 105-108.

1921 Gene, S. H. (2007). *Fish conservation: A guide to understanding and restoring global*
1922 *aquatic biodiversity and fishery resources.* Island press, USA.

1923 Goudet, J. 1995. Fstat (version 1.2): a computer software to calculate *F*-statistics. J.
1924 Hered. 86: 485 - 486

1925 Guo, S. W., E. A. Thompson. 1992. Performing the exact test of Hardy-Weinberg
1926 proportions for multiple alleles. Biometrics. 48: 361 - 372.

1927 Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis
1928 program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.

1929 Hare, M. P., L. Nunney, M. K. Schwartz, D. E. Ruzzante, M. Burford, R. S. Waples, K. Reugg
1930 and F. Palstra. 2011. Understanding and Estimating Effective Population Size for
1931 Practical Application in Marine Species Management. Conservation Biology 25(3):
1932 438 – 449

1933 Harrison, A.C. 1963. The Olifants/Doorn River system and its fishes. *Piscator,* 57**:** 25-57.
1934 Humphries, P.; King, A.J. and Koehn, J.D. 1999. Fish, flows and flood plains: links
1935 between freshwater fishes and their environment in the Murray-Darling River
1936 system, Australia. *Environmental Biology of Fishes,* 56**:** 129-151.

1937 Harrison, A.C. 1977. The early transactions of the Cape Piscatorial Society. *Piscator,* 98**:**
1938 118-123.

1939 Hubisz, M. J., D. Falush, M. Stephens, J. K. Pritchard. 2009. Inferring weak population
1940 structure with the assistance of sample group information. Mol. Ecol. Res. 9: 1322 –
1941 1332.

Impson ND, MS Marriott, IR Bills & Skelton PH. 2007. Conservation biology and management of a critically endangered cyprinid, the Twee River redfin, Barbus erubescens (Teleostei: Cyprinidae), of the Cape Floristic Region, South Africa African Journal of Aquatic Science 32(1): 27-33.

Impson, D. and Swartz, E. 2007. *Labeo seeberi In: IUCN 2010. IUCN Red List of Threatened Species. Version 2010.4* [Online]. Available: www.iucnredlist.org [Accessed 24 March 2011].

Impson N. D., Hamman K. C. D. 2000.Conservation status, threats and future prospects for the survival of freshwater fishes of the Western Cape Province, South Africa. In: Le Roux IG (ed.) Management and Ecology of River Fisheries. Blackwell, London. 418 – 427

Impson, N.D.; Bills, R.I. and Cambray, J.A. 2000. State of Biodiversity: Western Cape Province, South Africa: freshwater fishes.

Impson, N. D. 1997. Threatened fishes of the world: *Labeo seeberi* Gilchrist & Thompson, 1911 (Cyprinidae). Environ. Biol. Fishes. 54; 44

Impson, N.D. 1995. Distribution of fishes in the Oorlogskloof and Koebee Rivers. Unpublished Report. Cape Nature Conservation, Jonkershoek.

Ivanova N. V., T.S. Zemlak, R. H. Hanner, P. D. N. Hebert. 2007. Barcoding: Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes 7, 544–548

Jubb, R.A. 1967. *Freshwater fishes of southern Africa,* Cape Town, Balkema.

Kalinowski, S. 2005. HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes. 5. 187 - 189. 10.1111/j.1471-8286.2004.00845.x.

Kearse, M., Moir, R., Wilson, A. *et al*. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28(12):1647–1649. doi:10.1093/bioinformatics/bts199

Kim, Kyung Seok, and Thomas W. Sappington. 2013. "Microsatellite data analysis for population genetics." In Microsatellites, pp. 271-295. Humana Press, Totowa, NJ, doi: 10.1007/ 978-1-62703-389-3_19.

King, J.M.; Cambray, J.A. and Impson, N.D. 1998. Linked effects of dam-released floods and water temperature on spawning of the Clanwilliam yellowfish *Barbus capensis*. *Hydrobiologia,* 384**:** 245-265.

Kocher T. D., W. K. Thomas, A. Meyer, S. V. Edwards, S. Pääbo, F. X. Villablanca, A. C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. Proc. Nat. Acad. Sci. USA 86: pp. 6196-6200

Koehn, J.D. and Harrington, D.J. 2006. Environmental conditions and timing for the spawning of Murray cod (Maccullochella peelii peelii) and the endangered trout cod (M. macquariensis) in southeastern Australian rivers. *River Research and Applications,* 22**:** 327-342.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour*. 15(5):1179–1191. doi:10.1111/1755-0998.12387

Lacy, R. C. 1997. Importance of genetic variation to the viability of mammalian populations. - J. Mammal. 78: 320-335.

Lacy, R. C. and Lindenmayer, D. B. 1995. A simulation study of the impacts of population subdivision on the mountain brushtail possum, Trichosurus caninus Ogilby (Phalangeri dae: Marsupialia), in south-eastern Australia. II. Loss of genetic variation within and between subpopulations. - Biol. Conserv. 73: 131-142.

Lacy, R. C. 1989. Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. Zoo Biology8,111–123.

Lande, R. 1995. Mutation and conservation. - Conserv. Biol. 9: 782-791.

Landguth, E. L., Fedy, B. C., Oyler-McCance, S. J., Garey, A. L., Emel, S. L., Mumma, M., Wagner, H. H., Fortin, M., Cushman, S. A. 2012. Effects of sample size, numberof

markers, and allelic richness on the detection of spatial genetic pattern. Molecular Ecology Resources 12; 276-284.

Leberg, P. L. (1991). Effects of genetic variation on the growth of fish populations: Conservation Implications. Journal of Fish Biology37;193–195.

Leidy, R. A., & Moyle, P. B. (1998). Conservation status of the World's Fish Fauna: An overview. In P. L. Fiedler, & P. M. Kareiva (Eds.), *Conservation biology* (pp. 187‑227). Boston, MA: Springer.

Leprieur, F., Brosse, S., Garcia-Berthou, E., Oberdorff, T., Olden, J. D. & Townsend, C. R. (2009). Scientific uncertainty and the assessment of risks posed by non-native freshwater fishes. *FISH and FISHERIES* 10, 88-97.

Librado, P. and Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25**: 1451-1452.

Linder, H. P., S. D. Johnson, M. Kuhlmann, C. A. Matthee, R. Nyffeler and E. R. Swartz. 2010. Biotic diversity in southern African winter-rainfall region. Environ Sustain 2: 109-116

Love Stowell, S. M., Pinzone, C. A., and Martin, A. P. 2017. Overcoming barriers to active interventions for genetic diversity. Biodivers. Conserv. 26, 1753–1765. doi: 10.1007/s10531-017-1330-z

Lowenstein, J. H., T. W. Osmundson, S. Becker, R. Hanner and M. L. J. Stiassny. 2011. Incorporating DNA barcodes into a multi-year inventory of fishes of the hyper diverse Lower Congo River, with a multi-gene performance assessment of genus *Labeo* as a case study. Mitochondrial DNA 22(1): 52 – 70

Lubbe, A., Jordaan, M., Bragg, C., Paxton, B.R., Schumann, M., Van der Walt, R. & Ngobela, T. 2015. *Labeo seeberi. The IUCN Red List of Threatened Species 2015*: e.T11071A76889617. http://dx.doi.org/10.2305/IUCN.UK.2015-2.RLTS.T11071A76889617.en

Lynch M, R. Lande. 1998. The critical effective size for a genetically secure population. *Animal Conservation*, **1**, 70–72.

Markert, J. A., Champlin, D. M., Gutjahr‐Gobell, R., Grear, J. S., Kuhn, A., McGreevy, T. J., Nacci, D. E. (2010). Population genetic diversity and fitness in multiple environments. *BMC Evolutionary Biology*, *10*, 205. https://doi.org/10.1186/1471-2148-10-205.

Marr, S. M. 2012. An assessment of a proposal to eradicate non-native fish from priority rivers in the Cape Floristic Region, South Africa. African Journal of Aquatic Sciences 37: 141-142

Martinez, A. S., Willoughby, J. R., Christie, M. R. 2018. Genetic diversity in fishes is influenced by habitat type and life-history variation. *Ecology and Evolution* 8: 12022-12031.

Moritz, C. (1994). Defining evolutionarily significant units for conservation. Trends Ecol. Evol. 9, 373–375. doi: 10.1016/0169-5347(94)90057-4

Nei, M., Maruyama, T., Chakraborty, R. 1975 Bottleneck effect and genetic-variability in populations. *Evolution*, **29**, 1–10.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from small number of individuals. Genetics89: 583 - 590

Nel, J.L.; Belcher, A.; Impson, N.D.; Kotze, I.M.; Paxton, B.R.; L.Y, S. and Smith-Adao, L.B. 2006. Conservation assessment of freshwater biodiversity in the Olifants/Doorn Water Management Area: Final Report. *CSIR/NRE/ECO/ER/2006/0182/C*. Centre for Industrial and Scientific Research (CSIR), Stellenbosch. 128 pp.

Ostergaard S, Hansen MM, Loeschcke V, Nielsen EE (2003) Long-term temporal changes of genetic composition in brown trout (*Salmo trutta* L.) populations inhabiting an unstable environment. *Molecular Ecology*, **12**, 3123–3135.

Palsbøll, P. J., Bérubé, M., and Allendorf, F. W. 2007. Identification of management units using population genetic data. Trends Ecol. Evol. 22, 11–16. doi: 10.1016/j.tree.2006.09.003

Palstra, F. P. and D. E. Ruzzante. 2008. Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? Molecular Ecology 17: 3428 – 3447

Pauly, D., & Zeller, D. (2016). Catch reconstructions reveal that global marine fisheries catches arehigher than reported and declining. *Nature Communications*, *7*, 10244. https://doi.org/10.1038/ncomms10244.

Paxton BR, Ramollo P, Schumann M and Jordaan M. 2012. Draft Biodiversity Management Plan for Species for the Clanwilliam Sandfish (Labeo seeberi). The Freshwater Consulting Group (FCG), Northern Cape Department Nature and Environment Conservation and CapeNature.

Paxton BR, Clark BM & Brown CA. 2002. An assessment of the effects of habitat degradation and exotic fish species invasions on the distribution of three endemic cyprinids: Barbus capensis, Barbus serra and Labeo seeberi in the Olifants and Doring Rivers, Western Cape. Report prepared for the Department of Water Affairs and Forestry and Department of Agriculture. DWAF Report No. PB E000-00-1302.

Peakall, R., Smouse, P. E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update, *Bioinformatics*, 28(19): 2537–2539, https://doi.org/10.1093/bioinformatics/bts460

PGWC. 2007. Western Cape Olifants/Doring River irrigation study. Summary Report for the Provincial Government of the Western Cape: Department of Agriculture. *Prepared by ARCUS GIBB (Pty) Ltd as part of Contract No. 259-2000/2001: Western Cape Olifants Doring River Irrigation Study. PGWC Report No. 259/2004/01*.

Puechmaille, S. J. 2016. The program *STRUCTURE* does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Molecular Ecology Resources 16; 608-627

Poff, N.L.R.; Allan, J.D.; Bain, M.B.; Karr, J.R.; Prestegaard, K.L.; Richter, B.D.; Sparks, R.E. and Stromberg, J.C. 1997. The natural flow regime. *Bioscience,* 47**:** 769-784.

Polzin, T., Daneshmand. S. V. 2003. On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters* 31; 12–20

Potts, W.M.; Booth, A.J.; Hecht, T. and Andrew, T.G. 2005. Reproductive biology of a riverine cyprinid, *Labeo umbratus* (Teleostei: Cyprinidae), in small South African reservoirs. *African Journal of Aquatic Science,* 30**:** 147-155.

Pritchard, J. K., M. Stephens, P. Donnelly. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics 155: 945 – 959

Queller, D.C. and Goodnight, K.F. 1989. Estimating relatedness using genetic markers. Evolution, 43: 258-275. doi:10.1111/j.1558-5646.1989.tb04226.x

Ralls, K., Ballou, J. D., Dudash,M. R., Eldridge,M. D. B., Fenster, C. B., Lacy, R. C., et al. 2018. Call for a paradigmshift in the genetic management of fragmented populations. Conserv. Lett. 11, 1–6. doi: 10.1111/conl.12412 6

Ramoejane, M. 2016. Gnetic diversity, evolutionary relationships and conservation of southern African Labeo fishes in relation to water management. Doctoral dissertation. Grahamstown: Rhodes University.

Ramollo, P. P. M. Schumann and W. A. J. Pretorius 2012. Freshwater fish species of the Oorlogskkloof River, Northern Cape Province, South Africa. *Our Nature*, 10; 8-16.

Raymond, M., F, Rousset. 1995. Population genetics software for exact tests and ecumenicism. *J. Hered*. 86: 248 - 249.

Reed, D. H., & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, *17*, 230‐237. https://doi.org/10.1046/j.1523-1739.2003.01236.x.

Reid, G.M. 1982. The form, function and phylogenetic significance of the vomero-palatine organ in cyprinid fishes. *Journal of Natural History,* 16**:** 497-510.

Reid, G.M. 1985. *A revision of African species of Labeo (Pisces : Cyprinidae) and re-definition of the genus. Thesis Zoologicae Vol. 6,* Braunschweig, J. Cramer Verlag.

Robeldo, D., Palaiokostas, C., Bargelloni, L., Martinez, P., Houstan, R. 2018. Applications of genotyping by sequencing in aquaculture breeding genetics. *Reviews in Aquaculture* 10; 670-682.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Galtier, N. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, *515*, 261‐ 263. https://doi.org/10.1038/nature13685.

Roth, D. 1952. A remarkable bag of smallmouth bass from the Olifants River, Clanwilliam. *Piscator,* 6**:** 68.

Sahoo, L., Sahu, B. P., Das, S. P., Swain, S. K., Bej, D., Patel, A., Jayasankar, P., Das, P. 2014. Limited genetic differentiation in *Labeo rohita* (Hamilton 1822) populations as revealed by microsatellite markers. *Biochemical Systematics and Ecology* 57; 427-431.

Schmelle,r D. S., Merila, J. 2007. Demographic and genetic estimates of effective population and breeding size in the amphibian *Rana temporaria*. *Conservation Biology*, **21**, 142–151.

Scott, H.A. 1982. The Olifants System - unique habitat for rare Cape fishes. *Cape Conservation Series 2*. 14 pp.

Scribner, K. T., W. H. Lowe, E. Landguth, G. Luikart, D. M. Infante, G. E. Whelan & C. C. Muhlfeld. 2016. Applications of Genetic Data to Improve Management and Conservation of River Fishes and Their Habitats, Fisheries, 41:4, 174-188, DOI: 10.1080/03632415.2016.1150838

Singh, R. K., Lal, K. K., Mohindra, V., Punia, P., Sah, R. S., Kuar, R., Gupta, A., Das, R., Lakra, W. S.. Ayyappan, S. 2012. Genetic diversity of Indian Major Carp, Labeo calbasu (Hamilton, 1822) populations inferred from microsatellite loci. Biochemical Systematics and Ecology 44; 307-316.

Skelton PH, Swartz ER. 2011. Walking the tightrope: trends in African freshwater systematic ichthyology. Journal of Fish Biology 79: 1413-1435.

Skelton, P.H. 2001. A Complete Guide to the Freshwater Fishes of Southern Africa. Struik Publishers, Cape Town, South Africa.

Skelton, P.H. 1994. Diversity and distribution of freshwater fishes in east and southern Africa. Annals of the Royal Central African Museum 275: 95-131.

Skelton, P.H. 1987. South African Red Data Book – Fishes. South African National Scientific Programmes Report 137: 1–199.

Spielman, D. B. W. Brook, R. Frankham. 2004. Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States*, 101, 15261– 15264.

Stockwell, C. A., Hendry, A. P., & Kinnison, M. T. 2003. Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution*, *18*, 94‑101. https://doi.org/10.1016/S0169-5347(02)00044-7.

Swartz E.R. 2013. Conservation genetic management of *Austroglanis gilli*, A. barnardi and Barbus erubescens. In I. R., Bills and N. D., Impson (Red.), Conservation biology of endangered freshwater fishes – Linking conservation of endangered freshwater fishes with river conservation, focussing on the Cederberg. (pp. 79 – 93). Report to the Water Research Comission.

Swartz E.R., Skelton PH & Bloomer P. 2008. Molecular phylogenetics and evolution phylogeny and biogeography of the genus Pseudobarbus (Cyprinidae): Shedding light on the drainage history of rivers associated with the Cape Floristic Region. Molecular Phylogenetics and Evolution 51:75-84.

Swartz, E.R., Flemming, A.F. & Mouton, P.L.N. 2004. Contrasting genetic patterns and population histories in three threatened redfin species (Cyprinidae) from the Olifants River system, western South Africa. Journal of Fish Biology 64: 153-1167.

Tamura, K., G. Stecher, D. Peterson, A. Filipskiand and S. Kumar. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution: 30, 2725-2729.

Thomaz, A. T., Christie, M. R., & Knowles, L. L. (2016). The architecture of river networks can drive the evolutionary dynamics of aquatic populations. *Evolution*, *70*, 731‑739. https://doi.org/10.1111/evo.12883.

Tómasson, T.; Cambray, J.A. and Jackson, P.B.N. 1984. Reproductive biology of four large riverine fishes (Cyprinidae) in a man-made lake, Orange River, South Africa. *Hydrobiologia,* 112**:** 179-195.

Tweddle, D., R. Bills, E. R. Swartz, W. Coetzer, L. Da Costa, J Engelbrecht, J. A. Cambray, B. Marshall, N. D. Impson, P. H. Skelton, W. R. T. Darwall and K. G. Smith. 2009. The

status and distribution of freshwater fishes. In: Darwall WRT, Smith KG, Tweddle D, Skelton P (eds) The status and distributionof freshwater biodiversity in South Africa. IUCN, Gland, Switzerland and SAIAB (South African Institute for Aquatic Biodiversity), Information Press, Oxford, pp 21-37

Vandewoestijne, S., Schtickzelle, N., & Baguette, M. 2008. Positive correlation between genetic diversity and fitness in a large, well‐connected metapopulation. *BMC Biology*, *6*, 46. https://doi.org/10.1186/1741-7007-6-46.

Van Rensburg, K.J. 1966. Die vis van die Olifantsrivier (weskus) met spesiale verwysing na die geelvis (*Barbus capensis*) en die saagvin (*Barbus serra*). Investigational Report Cape Department of Nature Conservation 10: 1-14. Cape Department of Nature Conservation, Cape Town.

Vila C, A. K. Sundqvist, O. Flagstad *et al*. 2003. Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proceedings of the Royal Society of London Series B-Biology Sciences*, **270**, 91–97.

Vrijenhoek, R.C. 1998. Conservation genetics of freshwater fish. Journal of Fish Biology, 53: 394-412. doi:10.1111/j.1095-8649.1998.tb01039.x

Waples, R.S. and Do, C. 2010. Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evolutionary Applications, 3: 244-262. doi:10.1111/j.1752-4571.2009.00104.x

Waples, R.S. 1989. A generalized-approach for estimating effective population-size from temporal changes in allele frequency. Genetics, 121, 379–391.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2007) Effective population sizes and migration rates in fragmented populations of an endangered insect (*Coenagrion mercuriale*: Odonata). *Journal of Animal Ecology*, **76**, 790–800.

Westemeier RL, J. D. Brawn, S. A. Simpson *et al*. 1998. Tracking the long-term decline and recovery of an isolated population. *Science*, **282**, 1695–1698.

Weyl, O. L. F., B. Finlayson, N. D. Impson, D. J. Woodford and J. Steikjer. 2014. Threatened Endemic Fishes In South Africa's Cape Floristic Region: A new Beginning for the Rondegat River. Fisheries 39(6): 270-278

Weyl, O. L. F., B. R. Ellender, D. J. Woodford, M. S. Jordaan. 2013. Fish distributions in the Rondegat River, Cape Floristic Region, South Africa, and the immediate impact of rotenone treatment in an invaded reach. Afr J Aquat Sci 38: 201-209

Willi Y, J. Van Buskirk, A. A. Hoffmann. 2006. Limits to the adaptive potential of small populations. *Annual Review of Ecology Evolution and Systematics*, **37**, 433–458.

Willoughby, J. R., Sundaram, M., Wijayawardena, B. K., Kimble, S. J. A., Ji, Y., Fernandez, N. B., ⋯ DeWoody, J. A. 2015. The reduction of genetic diversity in threatened vertebrates and new recommendations regarding IUCN conservation rankings. *Biological Conservation*, *191*, 495‑503. https://doi.org/10.1016/j.biocon.2015.07.025.

Woodford, D.J.; Impson, N.D.; Day, J.A. and Bills, I.R. 2005. The predatory impact of invasive alien smallmouth bass, Micropterus dolomieu (Teleostei: Centrarchidae), on indigenous fishes in a Cape Floristic Region mountain stream. African Journal of Aquatic Science, 30(2): 167-173.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics*, **16**, 97–159.

Yang, L. and R. L. Mayden. 2009. Phylogenetic relationships, subdivision, and biogeography of cyprinid tribe Labeonini (*sensu* Rainboth, 1991) (Teleostei: Cypriniormes), with comments on the implications of lips and associated structures in the labeonin classification. Molecular Phylogenetics and Evolution 54: 254 – 265

Yang, L., M. Arunachalam, T. Sado, B. A. Levin, A. S. Golubtsov, J. Freyhof, J. P. Friel, W. Chen, M. V. Hirt, R. Manickam, M. K. Agnew, A. M. Simmons, K. Saitoh, M. Miya, R. L. Mayden and S. He. 2012. Molecular phylogeny of the cyprinid tribe Labeonini (Teleostei: Cypriniformes). Molecular Phylogenetics and Evolution 65: 362 – 379

Zheng, L., J. Yang and C. Chen. 2012. Phylogeny of the Labeoninae (Teleostei, Cypriniformes) based on nuclear DNA sequences and implications on character evolution and biogeography. Current Zoology 58(6): 837 – 850

Addendums:

Addendum A

Table: *CO1* Accesions downloaded for Genus Labeo tree construction

COI accessions	
Species	Accession
Labeo altivelis	CTOL3912
Labeo weeksii	CTOL00405
Labeo horie	CTOL3960
Labeo senegalensis	BNF 171
Labeo barbatus	t-062-6196
Labeo batesii	t-067-6671
Labeo lineatus	t-075-7407
Labeo seeberi	This study
Labeo vulgaris	CTOL3966
Labeo forskalii	CTOL3959
Labeo nasus	t-071-7005
Labeo simpsoni	t-074-7372
Labeo annectens	t-062-6119
Labeo parvus	CTOL3957
Labeo lukulae	CTOL00402
Labeo quadribarbis	t-064-6380
Labeo coubie	BNF 159
Labeo longipinnis	CTOL3952
Labeo bata	NF731
Labeo boga	NF240
Labeo boggut	NF636
Labeo dyocheilus	NF251
Labeo pierrei	N/A
Labeo yunnanensis	CTOL3956
Labeo calbasu	NF239
Labeo chrysophekadion	N/A
Labeo barbatulus	N/A
Labeo fimbriatus	LF-1019
Labeo rohita	PR009205
Labeo stolizkae	N3
Labeo caeruleus	NF692
Labeo gonius	WL-F82
Labeo dussumieri	LD-1
Labeo rajasthanicus	LR 2

Addendum B

Table: *Cytb* Accesions downloaded for Genus Labeo tree construction

Cytb accessions	
Species	Accession
Labeo altivelis	CTOL3912
Labeo barbatulus	N/A
Labeo bata	CTOL01920
Labeo boggut	NBFGRLBG-149
Labeo calbasu	NBFGRLK-1083
Labeo chrysophekadion	N/A
Labeo coubie	CTOL3163
Labeo dussumieri	NBFGRLD-4
Labeo dyocheilus	CTOL01922
Labeo fimbriatus	LF531
Labeo forskalii	T10
Labeo gonius	NBFGRLG-221
Labeo horie	CTOL3960
Labeo longipinnis	CTOL3952
Labeo lukulae	N/A
Labeo parvus	CTOL3957
Labeo pierrei	N/A
Labeo rohita	LR.NAB.2239
Labeo sorex	AMNH233629
Labeo stolizkae	KIZCXY20060059
Labeo vulgaris	CTOL3966
Labeo weeksii	N/A
Labeo yunnanensis	N/A

Addendum C

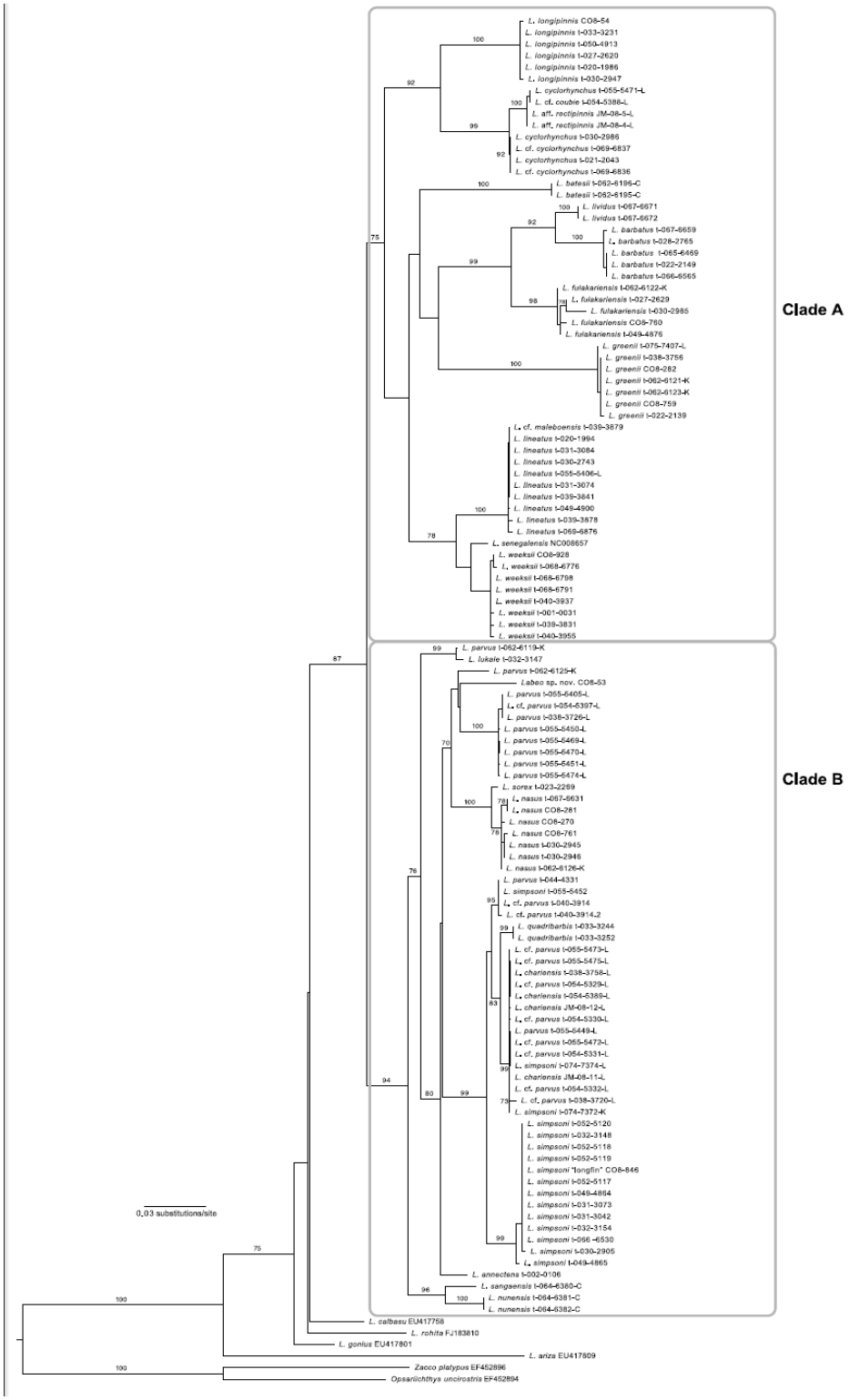


Figure: *CO1 Labeo* gene tree. Maximum-likelihood phylogram showing branch bootstrap support above the 70% (log-likelihood score of best tree: 24137.464790). Outgroup taxa omitted. Circle, Asian *Labeo*; "a" and "b", the two main clades discussed in the text. Specimen codes ending in "K" denote individuals caught in the vicinity of Kisangani (Upper Congo); "L" collected from the Lualua, a large southern tributary of the Congo River; and "C" describes specimens collected from Cameroon. One species, *L. senegalensis* (GenBank accession number NC008657), inhabits western Africa.

Addendum D

Table: List of primers used for microsatellite analysis including each primers nucleotide sequence, accession number, expected and observed heterozygosity, repeat type and reference

Locus Name	Primer sequence	Accession no.	Repeat type	No. alleles	He	Ho	reference
Lr-36	F: 5'-AGC GTG TCT GAT GTG TGA AAG G-3' R: 5'-TCA GAT GCC TCC TGC ATT CTG-3'	AM269526	(CA)10	3	0.642	0.833	S. SWAIN1, S. P. DAS1, Evaluation of genetic variation in *Labeo fimbriatus* (Bloch, 1795) populations using heterologous primers *Indian J. Fish., 60(1)* : *29-35, 2013*
Lro-26	F: AGA TCA TTG CTG GGG AGT GTT TAT R: GAC CTG CCT GTG CCA TCT GTA	AM184144	(GT)29	3	0.649	0.520	
Lr-30	F: ACG CGC TAG GGT CGT ACA GTG R: CAG CAT CAT GTT AAG CGC TGTC	AM231179	(AC)15	3	0.631	0.846	
Lr-28	F: TTC ACG GAC AGA TTT GAC CCA G R: AGT CTT TTC AGG AGA TTA GCA G	AM231177	(AC)18	8	0.826	0.833	A. Patel, P. Das, S. K. Swain Development of 21 new microsatellite markers in Labeo rohita (rohu) Animal Genetics, 40, 251–254
Lr-29	F: ACG TAA AGG TCA CAA GCT GAA G R: AGC ACG GTG TTT GTG TGC GAG	AM231178	(AC)10	5	0.820	0.722	
Lr-46	F: TGA CGT ATT GTC AAC TAT GGT G R: TCC ACC TTC AAT ACC ATG ACT G	AM269536	(CA)18	4	0.757	0.764	
LF-4	F: GGC CAG TGT GAC ACA AAC A R: GTC CCG GAG TCT AAA GAC GAA C	JQ838159	(CT)12	9	0.743	0.607	S. Swain • S. P. Das Isolation and characterization of sixteen microsatellite loci for fringe-lipped carp, Labeo fimbriatus Conservation Genet Resour (2012) 4:913–915
LF-8	F: GTG AAG CAA CGA CTT CAG AGA G R: CCA GAA GAC CAT AGC AAC CAC	JQ838163	(GT)6	5	0.765	0.833	
LF-15	F: ACA CTC ACA CTC GCT CAC TCA C R: CGG TGA ATG CTG ATG AAC TG	JQ838170	(TC)10	6	0.803	0.806	
LF-16	F: AAC GTC ACA CAT GCT CCT AGT C R: CTG CCC ATG ACA CTG AAA CTC	JQ838171	(GA)26	5	0.724	0.781	

Lr-41	F TCC AGT CAC CAC ATG CGT TTG R GTC GAT TTC ATC GTG AGG CTC	AM269531	(GT)16		0.823	0.706	A PATEL1, P DAS2, **Test of Mendelian segregation and linkage relationships among 69 microsatellite loci in rohu (*Labeo rohita*)** *Indian Journal of Animal Sciences* **81** (8): 128–00, August 2011
Lr-44	F CAC CCA GGG AGT TAG TTT CTG R AAA GAG CAT CAT GGC ATT GAC	AM269534	(CA)25		0.844	0.801	

Addendum E

Table: Genetic diversity estimates for the 6 microsatellite loci across the three populations of *Labeo seeeri.*

OKNR	N_{ind}	A_N	A_R	A_{PR}	HWE	H_{Enb}	H_O	F_{IS}	Fr_{null}
Lro_26	82	8	3,6634	0,8891	0.1039	0.5885	0.7073	-0.2034	0.0000
Lr_36	82	7	5,2163	0,0851	0.6904	0.6998	0.6951	0.0068	0.0000
Lr_41	82	20	10,0657	2,4573	0.0000 *	0.9001	0.5244	0.4189	0.2185
LF_8	81	2	1,9999	0,0000	0.2157	0.4513	0.3827	0.1527	0.3849
LF_15	81	23	11,3338	3,0594	0.0004 *	0.9235	0.8642	0.0646	0.0342
LF_16	82	7	4,1322	0,4705	0.0818	0.6270	0.6829	-0.0899	0.0614
Mean	-	11,1667	6,0686	1,1602	0,1820	0,6984	0,6428	0,0583	0,1165
Riet									
Lro_26	36	4	3,1106	0,7957	0.6341	0.5442	0.4722	0.1339	0.0298
Lr_36	36	8	5,8255	0,7626	0.2964	0.7735	0.7500	0.0308	0.0242
Lr_41	36	17	11,0879	2,7113	0.0000 *	0.9206	0.4722	0.4906	0.2576
LF_8	36	2	1,9999	0,0000	0.7084	0.4507	0.5000	-0.1111	0.2887
LF_15	36	21	11,9832	3,3651	0.3099	0.9323	0.9167	0.0170	0.0111
LF_16	36	6	3,6662	0,1437	0.9189	0.5759	0.6389	-0.1111	0.0000
Mean	-	9,6667	6,2789	1,2964	0,4780	0,6995	0,6250	0,0750	0,1019
Bos									
Lro_26	10	4	4	1,5406	0.5437	0.5737	0.7000	-0.2353	0.0000
Lr_36	10	6	6	0,3953	0.1586	0.7211	0.7000	0.0308	0.0000
Lr_41	10	9	9	1,7005	0.0000 *	0.8737	0.3000	0.6687	0.2926
LF_8	10	2	2	0,0000	1.0000	0.2684	0.3000	-0.1250	0.0000
LF_15	10	13	13	4,3653	1.0000	0.9632	1.0000	-0.0405	0.0000
LF_16	10	5	5	0,8341	0.5062	0.6526	0.5000	0.2437	0.1033
Mean	-	6,5000	6,5000	1,4726	0,5348	0,6755	0,5833	0,0904	0,0660

* = significantly different ($P < 0.05$) from HWE. N_{ind} = Number of individuals; A_N = Number of observed alleles; A_R = Allelic richness; A_{PR} = Private allelic richness HWE = Hardy Weinberg Equilibrium; H_{Enb} = Nei's unbiased expected heterozygosity; H_O = Observed heterozygosity; F_{IS} = Inbreeding coefficient;
Fr_{null} = Null allele frequency.

Addendum F

Table: All pops F_{ST} *P*-values

	OKNR	Riet	Bos
Riet	0.04395+-0.0071	*	
Bos	0.44141+-0.0177	0.35059+-0.0162	*

Addendum G

Table: OKNR F_{ST} *P*-values

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8	Site 9	Site 10	Site 11
Site 2	0.23242+-0.0120	*									
Site 3	0.56543+-0.0164	0.08496+-0.0081	*								
Site 4	0.93164+-0.0070	0.34961+-0.0133	0.59961+-0.0133	*							
Site 5	0.50195+-0.0191	0.08203+-0.0090	0.50195+-0.0175	0.82031+-0.0111	*						
Site 6	0.57812+-0.0145	0.17480+-0.0149	0.02637+-0.0039	0.72559+-0.0124	0.54004+-0.0123	*					
Site 7	0.97363+-0.0040	0.10742+-0.0116	0.51660+-0.0195	0.90039+-0.0093	0.67969+-0.0128	0.70703+-0.0150	*				
Site 8	0.72949+-0.0106	0.32715+-0.0131	0.28613+-0.0126	0.98438+-0.0031	0.50781+-0.0138	0.68848+-0.0137	0.81250+-0.0125	*			
Site 9	0.32129+-0.0139	0.28711+-0.0128	0.50098+-0.0149	0.41113+-0.0160	0.83691+-0.0089	0.15723+-0.0117	0.45020+-0.0147	0.19922+-0.0118	*		
Site 10	0.56641+-0.0141	0.27344+-0.0124	0.16016+-0.0099	0.58301+-0.0135	0.57129+-0.0138	0.70605+-0.0120	0.62500+-0.0145	0.23145+-0.0156	0.33691+-0.0140	*	
Site 11	0.02246+-0.0040	0.32910+-0.0166	0.02539+-0.0042	0.34668+-0.0162	0.13672+-0.0098	0.02246+-0.0047	0.04688+-0.0067	0.02930+-0.0041	0.06445+-0.0078	0.22754+-0.0126	0.00000

Addendum H

Table: All pops ϕ_{ST} *P*-values

	OKNR	Riet	Bos
Riet	0.22754	0.00000	
Bos	0.24609	0.06543	0.00000